JAN 22 2007

617341759

QA
76.7
D3
D483
2006

"SO, YOU WANT A DATABASE?:
A BEGINNER'S GUIDE
TO DATABASE TECHNOLOGY
FOR SMALL CULTURAL INSTITUTIONS"

By

Shannon Leigh Perry
Bachelor of Arts,
University of Guelph
2002

A Thesis Project
Presented to
Ryerson University
and
Gorge Eastman House,
International Museum
of
Photography and Film

In Partial Fulfillment
of the
Requirements for
the Degree of
Master of Arts
in the Program of

Photographic Preservation
and
Collections Management

Toronto, Ontario, Canada, ~~2005~~

©Shannon Perry 2006

PROPERTY OF
RYERSON UNIVERSITY LIBRARY

UMI Number: EC53618

INFORMATION TO USERS

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleed-through, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

UMI[®]

UMI Microform EC53618
Copyright 2009 by ProQuest LLC
All rights reserved. This microform edition is protected against
unauthorized copying under Title 17, United States Code.

ProQuest LLC
789 East Eisenhower Parkway
P.O. Box 1346
Ann Arbor, MI 48106-1346

I, Shannon Perry, hereby declare that I am the sole author of this thesis or dissertation. I authorize Ryerson University and George Eastman House International Museum of Photography and Film to lend this thesis or dissertation to other institutions or individuals for the purpose of scholarly research.

I, Shannon Perry, further authorize Ryerson University and George Eastman House International Museum of Photography and Film to reproduce this thesis or dissertation by photocopying or by other means, in total or in part, at the request of other institutions or individuals for the purpose of scholarly research.

"So, You Want a Database?: A Beginners Guide to Database Technology for Small Cultural Institutions."

Master of Arts, 2006

Shannon Leigh Perry, Photographic Preservation and Collections Management, Ryerson University and George Eastman House, International Museum of Photography and Film.

This thesis project is comprised of a twenty page folded booklet and ten page analytical paper. The booklet is meant as a beginner's guide to collections management systems within cultural institutions. Its primary audience is smaller institutions that may not have the benefit of staff with expert or prior knowledge of electronic databases and collections management systems. The booklet outlines what a database is, why collection management systems are important, a frequently asked question section, a glossary and an additional resource page. The frequently asked questions section gives examples of why database management systems are important and how to choose a collections management system, along with other commonly asked questions.

Acknowledgments:

I would like to express my gratitude to all the George Eastman House and Ryerson staff that helped me realize my topic and finish my project. I would like to especially thank my thesis advisor Dr. Alison Nordstrom for her excellent guidance and patience. I would also like to acknowledge the numerous people with whom I have communicated over the last year and discussed this thesis project. Thank You!

Table of Contents

Introduction	1
Developing the Idea	3
Identifying a need	5
Compiling and Evaluating Resources	7
Example of Literary Review	9
1. Every Museum needs a database	10
2. You need to develop a budget and an IT plan for implementing the system	12
3. Every museum needs a commitment by the staff to use the Software	13
4. You will need a designated database administrator to serve as “data cop”	14
5. Regularly back up your data	14
Conclusion of Literary Review	15
Overview of Booklet	16
Title	17
Content	18
1. General Information	18
2. Frequently Asked Questions	19
3. Glossary	19
Conclusion/ Anticipated Use of the Booklet	21
Bibliography	22
Copy of Booklet	(pocket) 25

Introduction

The purpose of this thesis is to help workers within small cultural institutions understand the basics of databases and collection management systems as they directly relate to cultural institutions and serve as a catalyst for future studies and projects. The intent of this project is not to evaluate the different database models and collection management systems available, but rather to address the common issues that novice users of collection management systems have repeatedly addressed in articles, conference forums, conversation and blogs. The institutions that would benefit most from this project are those that do not have an IT staff member or department and do not have the time or money to send existing staff members for additional training.

I recognized this to be an important subject area that needed to be addressed by professionals in the field. Perhaps the reason it has not been addressed is because most museums already have some type of electronic database and collection management system. In small cultural institutions this not always the case and their pleas for assistance are often unheard.

Small cultural institutions are operating with minimal staff and operating budgets, and are often located in the more rural areas of the country. They may not have the access to resources available to larger institutions, such as staff trained in electronic records management or even those with a formal education in museum studies. A majority of the staff may be volunteers, and as I heard from one cultural worker, if the staff has a computer system and something goes

wrong, they are often relying on a family member with more technical experience to fix things rather than a trained IT professional. I feel that cultural workers at larger institutions take for granted the fact that if they have a problem understanding any aspect of the collection management program, there is usually another staff member they can talk to. This is not usually possible in smaller institutions, which is why I created this project – to share information that I have had the benefit of learning both as part of my Masters degree and as part of working within several large cultural institutions.

Developing the Idea

The concept for my practical project began last summer during my internship at the Canadian Museum of Contemporary Photography (part of the National Gallery of Canada). The project I was given was to conduct a feasibility study for creating electronic access to a collection of a few thousand negatives, transparencies and copy prints, which at the time, were not entered into any type of searchable database. The original idea for the project was to create a separate database due to the unique nature of the collection – perhaps build on the work previously completed in past years.

Before completing the feasibility study I researched all available material on electronic access in museum collections. I researched collection management systems, digitization standards and project guidelines. In addition to the printed material I was able to find (totaling well over six hundred pages), I also had access to the staff of the National Gallery, including members of the Information Technology department, who openly discussed their experiences with the collection management system used by the National Gallery – a Willoughby product called Multi MIMSY.

Throughout my internship project I kept wondering “What would I be doing if I didn’t have the benefit of knowledgeable staff with whom I could discuss the specifics of database technology and management procedures?”. I directly benefited from working in a large institution, but what about the small cultural institutions – institutions that are operating with only a few staff members or

volunteers? I felt that my thesis project would be an excellent opportunity to share what I had learned and experienced during my internship.

The original concept for this thesis project was meant to address how cultural institutions could get more use out of their existing databases, but after discussion with my thesis advisor and several other members of the museum community, it became clear that it would perhaps be more beneficial to start at the beginning – what *is* a database exactly and why should we – as protectors of our culture’s heritage and history – use them.

Identifying a Need

My first museum related job was at a small museum run mostly by volunteers. Although the museum is small in terms of staff, budget and square footage - the collection of military vehicles and every other type of war related ephemera imaginable has the potential to be of interest to people across Canada. The information regarding the collection was being entered into various Microsoft Word documents and Excel spreadsheets. There was no searchable database, and the information that was being entered into an electronic record was inconsistent at best.

Last year I spent quite a bit of time at the local archives in Whitby. The archivist did not have a computer, let alone a searchable database. The collection information was stored in his head – and since he was the sole employee at the archives, I do not need to explain why this is a potentially dangerous situation in regards to understanding the content of the collection. In speaking with several members of the museum community over the past year, I heard repeatedly that situations such as these are far from uncommon.

In 1999 a survey was conducted by the Canadian Heritage Information Network (CHIN) which established the level of computer use in cultural institutions of all sizes across Canada. The findings, which were published in 2001 under the title “Information Technology in Canadian Museums: A Survey by

the Canadian Heritage Information Network”¹, lend further credence to my own experiences with small cultural institutions. The survey indicated that as of 1999, 25% of the community still did not have *any* form of computer technology. For the institutions that *did* have computer technology, the most extensive use was within the collection management area. However, much of the collection information that was catalogued was still not in electronic form. Only 40% of the institutions using computers had a Collections Management System (CMS) or Database Management System (DBMS).

Although the survey is from 1999, I feel that it is still an accurate gauge of the situation in smaller cultural intuitions in Canada². Even if half of the institution that did not have computer systems or CMS/DBMS in 1999 have one today, there are still a fair number of institutions in need of electronic database and database management – the majority being small institutions.

¹ Information Technology in Canadian Museums: A Survey by the Canadian Heritage Information Network (2001) http://www.chin.gc.ca/English/Reference_Library/Information_Technology/index.html

² 32.5% of returns, a majority, were from Community Museums, 67% classify their museum as not-for-profit, 30% of the institutions operate only in a particular season, while 57% operate year round, 85% of the institutions have 0-10 full-time staff, almost 45% have 0-10 volunteers working at the museum.

Compiling and Evaluating Resources

I conducted the majority of my research through on-line searches, and interviews with museum professionals in a variety of positions and institutions. I also searched through museum related magazines, journals and list-serves.

I chose to start my research using mostly information found on-line because I felt that this is where a cultural worker in a small institution who is looking for database and collection management system information would most likely start, since access to the internet is much more easily attainable than access to a large university or museum library collection. Also, the smaller institutions may not be aware of, or able to afford the cultural based journals and/or memberships (Museum Computer Network for example) which allow access to on-line articles, list serves and chat boards.

I was interested to see what kind of information was currently available – especially to someone who may not be aware of what key words to search for. By entering the words “database” in Google, for example, I was interested to see if the information was relevant for cultural institutions, what level of technical knowledge is required to fully understand the information provided, and who was providing the information.

Most of the information I found for this example can be placed into one of three general categories. It is either from software companies advertising their systems, heavily technical explanations and proposals for database operations published by IT professionals, or information related to business applications.

Searching by “museum database” did not offer much help either, as most of the results were for museums with an on-line database. By beginning my search in this manner I was able to see how people unfamiliar with the topic could become quickly frustrated. I already knew a fair amount about the subject, so I was able to scan an item quickly and determine whether or not it contained information relevant to the topic. The cultural workers for whom I am creating this project however, do not have the benefit of prior knowledge, and a lot of time could be misspent reading through information that does more to confuse than help.

It is important to look at who is publishing the information, because with on-line resources, there are no standards or quality control, and therefore, there is room for misinformation and error. There is also the possibility that the material is from a commercial source, in which case there is great potential for bias regarding a particular product, service or method.

I discovered that a majority of the information relating to collection management systems *and* cultural institutions is for people who already have a firm grasp of the technology and management principles involved. There is an overwhelming amount of information regarding how to ‘digitize’ a museums collection, but very little written detailing why, or more importantly – what to do *before* thinking about digitization.

I found resources that explained from a very basic level the principles of database design³, but they were written for users with an interest in business, rather than cultural institutions. I also searched for definitions of the key terms I

³ Database Demystified by Andrew J. Oppell, McGraw-Hill/Osborne, Emeryville, CA, 2004, How Software Works, by Ron White, Ziff Davis Press, Emeryville, CA, 1995.

was continually coming in contact with in my research, but never found a single source in which all of the applicable terms were explained in a language that was easily understandable for non-experts or applicable to how databases and collection management systems are used in cultural institutions.

Essentially, I was finding bits and pieces that I felt would be useful for small cultural institutions, but there was no single document related to this topic that could quickly outline the important issues, resources and terms available. Since most small cultural institutions do not have the time and/or resources to search for information, evaluate its content and distill the findings, I was inspired to create a small booklet that could be consulted easily to answer questions regarding terminology and the use of collection management systems, and begin to point people in the right direction.

Example of Literary Review

During my research I came across a published presentation given by a software company at a museum conference (Museum Computer Network) in 2002 entitled "Universals of Collection Management System Implementation"⁴. It was written and published by Willoughby Press, the company that licenses the MultiMimsy Collection management system, among other products. I was instantly intrigued by the title, but hesitant on the content considering the source. I decided to evaluate the content based on the research I had conducted and see if the information was biased to the Willoughby system, or if there were valuable

⁴ Universals of Collection Management System Implementation - Willoughby press 2002 - Delivered at Museum Computer Network Conference 2002 <http://www.willo.com/services/articles.asp>

points of information that could be extracted from the presentation. This report outlined five universals of collection management system implementation.

1. Every Museum needs an electronic database.

According to the presentation, relational databases allow you to keep detailed data about all kinds of elements related to your collection. It is important to maintain records of all the things that add to the history and life of the objects in the collection such as loans and exhibits. While I do not disagree with the argument, I do take issue with the blanket statement “every museum needs an electronic database”. Part of the reason I undertook this project was because I feel that too many institutions believe this and are acquiring an electronic database before giving full consideration to how they will use it and *why* they should be using it.

According to the presentation, until about 15 years ago or so, museums that pursued computerized collections management options had to build their own software. Oracle appeared in 1980 and object-based collection management systems appeared in the 1990's ⁵. The presentation goes on to state that “while admirable at the time, the programmers quit, retired or died, leaving dismal documentation and a small squadron of museum professionals who had no idea how to use the program/software”. While this paints a very gloomy picture indeed

⁵ Dr. Howard Besser web site <http://besser.tsoa.nyu.edu/howard/http://math.hws.edu/vaughn/cpsc/343/2003/history.html>

Chapter 6 from “Funding a Revolution: Government Support for Computing Research, National Academy Press, Washington D.C., 1999 <http://www.nap.edu/readingroom/books/far/sh6.html>

Database Demystified by Andrew J. Oppell, McGraw-Hill/Osborne, Emeryville, CA, 2004, How Software Works, by Ron White, Ziff Davis Press, Emeryville, CA, 1995.

of people who tried to create an in-house system, some brilliant innovations often come out of experimenting with, adapting or otherwise challenging the collections management systems available at the time.

According to the presentation, next came the out-of-the-box solutions that were offered by established companies (such as Willoughby) that work exclusively with museums, historical societies, and archives, and are the best option to avoid these pitfalls.

“You are provided with documentation and technical support, so even if your staff turns over, you can be confident that your questions will be answered by the vendor. Also, when you purchase out-of-the-box software, you will receive upgrades to the product, hopefully on a regular schedule. If you are tied to an in-house system, you are at the mercy of the programmer, or you may never get an update to your system to bring it into the twenty-first century”.

Although the presentation gives a good sell, the “out-of-the box” system is also not without faults. What is not mentioned is that companies go out of business or are bought and incorporated into bigger companies. Often technical support can be inconsistent - especially if the head offices or even support management are located in another country, and missing an upgrade can leave your system outmoded within a few years since system upgrades can happen every few months. Repeatedly I have read or heard from people within the IT field that upgrades are the quickest path to the death of a system. Once you commit yourself to a collection management system, you need to stay on top of all the system upgrades. What is often not told during the sales pitch is that support and upgrades can become expensive.

What was not mentioned in the presentation, but is a definite benefit from having a collection management system that caters specifically to cultural

institutions, is that you may be on the same system as another institution, making it easier to collaborate on joint projects on-line.

2. You need to develop a budget and an IT plan for implementing the system.

This *is* true, and it is a big factor when considering what type of collection management system your institution should get or if you already have one, to protect your previous investment. I found this point to be true regardless of the system used. What could have been included in this section is more explanation of many of the terms used such as network, server, data conversion, etc. – all of which I decided to include in the glossary, for if the terms are used in this type of article without any explanation, it is possible that people will lose interest in the topic if they have to look elsewhere continually for definitions and understanding.

According to the presentation, implementation costs do not stop once the program has been installed. Included in the budget should be plans for upgrades of both software and computers, training of staff, buying additional licenses, paying for annual support, and if you are switching collection management systems, possibly data conversion. If possible, save some extra funds in the budget for memory and operating system upgrades, because you will want to ensure that there is enough memory space for the information and that the records are being saved in a safe and secure system.

In the age of ever increasing budget cuts, it may be unrealistic to budget for the inevitable changes in technology, but you should at least make the staff at

your institution aware of the possible costs that may be incurred in a few years time and plan accordingly.

3. Every museum needs a commitment by the staff to use the software.

With this point I wholeheartedly agree. Even if a museum purchases a small collections management software tool, it will be investing money in implementation and training. If the staff is sufficiently trained in the database management software program, then, ideally, the program should help *any* museum professional work smarter and more efficiently while not demanding any changes to museum procedure.

The presentation states that if the company you are getting the software from offers training, you should be sure to use it. This point shows that the presentation is not particular about what company you use, only that you may want to look into a product that offers staff tutorial. Sometime – usually with the bigger, more complicated systems, the company will send an actual staff member to the institution to teach the museum staff. Other times the training is conducted through on-line tutorial or phone calls. Either way, it is important to understand and secure the type of training offered for the institution's staff.

In reality, what happens all too often is that a few people learn how to work in and/or navigate the collection management system, and the other staff members learn only the parts of the system that they feel may be relevant to their job or department. If they need to find something that reaches outside of this

scope, often they will simply ask a more knowledgeable staff member to retrieve the information, rather than trying to learn for themselves.

On the other side of the argument, some institutions feel that employees working within the collection management system should know only the fields in which they are working daily, to prevent people from having access to, or changing records that they shouldn't – such as insurance information.

4. You will need a designated database administrator to serve as “data cop”

This person is the only one who can perform global data changes, customize the system and work with creating, changing, and deleting user profiles. Data standardization is KEY to making your system work.

I agree with the point, but feel that the presentation could have expanded into why data standardization is key and how to ensure the information follows standards already agreed upon by the larger museum community. I felt that there was room for more information, but perhaps the authors felt it would be better suited for another article or presentation.

5. Regularly back up your data.

Regardless of how many people will be using the system, or how many computers will be used in the museum, every institution should invest in a dedicated server, which is a powerful computer (usually kept offsite) that is dedicated to running your collections management software, to storing your data, to keeping your collection's multimedia files in addition to the computers that are

used in the day-to-day management of the database. This is important so that incase something happens to the computers in your institution or office, you will have a back-up of the system safe and secure in another location. It also ensures that the information stays clean of any viruses that may be in an individual computer. It is important to make sure that the information within the server is updated regularly. If it is not possible to have a separate dedicated server, a small institution can regularly save the information to a series of disks or other recordable medium and store the saved information in another location – incase there is ever an accident at the institution such as theft of the computers, fire or floods.

Conclusion

In this presentation I was able to see that although there was some information lacking, such as terminology, there were some key points which could definitely be beneficial for smaller institutions to know about and which I could incorporate in part in this booklet. This presentation is also a good resource to include because it was presented at the Museum Computer Network in 2002, and as mentioned earlier, most small cultural institutions do not have the resources (time, money or staff) to attend such conferences and therefore may benefit from such information sessions.

Overview of Booklet

I am creating this booklet so that it can be an easy reference for cultural workers when they are working with databases and collection management systems. I want to create something compact and easily accessible, which is why I am choosing to create an actual paper based booklet instead of a web article or on-line tutorial. I feel that if people need to learn the basics of databases and collection management systems and related terminology, chances are, if given the option, they will feel more comfortable looking things up in a book then searching on-line.

I decided to create my project in a booklet form of this size (8 ½" x 11" sheets of paper – folded into booklets 8 ½" x 5 ½") rather than full size sheets bound in a different manner because I felt that the smaller size lent well to the concept of reference and flipping pages back and forth – such as flipping to the back for the glossary. This was confirmed when I showed the mock-up booklet to several museum workers in small cultural institutions⁶. Both people held the booklet and flipped through it, going back and forth through the pages. As we learned in our publishing class with Douglas Hollely, no matter how quickly you can access the material, people still want to hold something physically when they are reading; the importance of touch cannot be underestimated.

I realized that there would need to be a limit as to how in depth I would explain certain areas of electronic records management, because if someone is

⁶ Brian Winters, Archivist at Whitby Archives and Bev Dietrich, curator at Guelph Museums.

new to a subject it is easier for him or her to develop a true understanding by building on simple concepts and gradually adding more complex ideas and information after a solid foundation has been laid.

In talking with my thesis advisor, I came to realize the importance of including images in the booklet. As was discussed earlier in the semester, readers like to see the objects you are discussing in your writing. The problem I encountered when looking for images that could be included is that since this topic is heavily situated within the computer technology and business worlds, most or all of the images I found did not directly relate to cultural institutions. The closest I could find were database models that described the schema of book based databases for libraries. In the end what I had to do was create diagrams from scratch, making sure to create visuals that cultural workers could relate to by using objects and fields that they are familiar with. By doing so, I was able to use figures that do not need to be referenced or cleared for copyright if I wish to publish and distribute this booklet in the future.

Title

Knowing that the main component of my thesis is a booklet that I hope will be distributed amongst the museum community, I knew that choosing the title was extremely important. I needed to think of a title that had a nice hook to it - to ensure that people would be intrigued and at least pick it up for closer examination – but I also had to make sure that it was not misleading or inappropriate. I went through several different titles – cross referencing along the

way with library records and search engines such as Google to make sure that the title was not already in existence. I finally settled on “So You Want a Database?: A Beginners Guide to Database Technology for Small Cultural Institutions” because I feel that simply put, it identifies both the potential user of this booklet as well as what the booklet contains in terms of subject matter.

Content

Within the booklet I divided the material into four major sections: general information, frequently asked question, a glossary and additional resources.

1. General Information

Within the general information section I have included: Introduction, What is a Database?, and Why use a Collection Management System?. In the introduction I explained what the point of the booklet is, and what I hope will happen when people read it. I also outline what the booklet does not contain so that in case anyone did not read the 'beginner' section of the title, they will clearly understand that this booklet is not meant for the advanced collection management system user, although some aspects such as the glossary and additional resource section may be useful to any level of user. Within the general section I also explain what a database and database management system are, and what the difference is. I explain why collection management systems are important and what can be done with a database. I included these sections so that the users of this booklet can better understand the basic principles of

database design, which will complement the frequently asked questions section and the glossary.

2. Frequently Asked Questions

For the frequently asked questions section, I included questions that I myself asked when first beginning this project, as well as question or variants of questions that I saw repeatedly on various cultural based listserves⁷. I also incorporated research I had conducted on various topics. This proved to be the most challenging section because it was difficult to determine the limits of my subject area. I decided not to include information regarding digitization or on-line collection other than the brief sections I have included. Digitization and on-line use of databases is a topic which could be incorporated into another set of booklets if this project works out well.

3. Glossary

If I found the frequently asked questions section challenging, this section was by far the most difficult to organize and write. First I had to identify key terms to include in the glossary. I began working on the glossary before I had written the actual text for the booklet because at the time, I felt that this was a simple and straight-forward section I could work on. Once I began writing the text section I realized that I had forgotten to include words, or there were sections in

⁷ Museum Computer Network Listserve mcn-l@mcn.edu
Archives & Archivist Listserve archives@listserv.muohio.edu
Visual Materials Listserve visualmat@uclink4.berkeley.edu

the text that could best be elaborated in the glossary section because it caused certain sections of the booklet to flow less well when I included lengthy explanations.

Writing the definitions for the terms was interesting because, for any given term, I consulted four or five sources. I used encyclopedias and websites, previously published technical guides and projects, as well as information technology and computer science text books. I would read all of the definitions and using as simple language as possible write an explanation in my own words. Sometimes the definitions were so similar across four examples, that the changes I made were not very big. Other times the definitions differed hugely depending on the intended audience of the original text, and in these instances I had to think about how to describe the term in simple terms specifically for someone working in cultural institutions.

Conclusion/ Anticipated Use of the Booklet

I foresee the booklet being used by cultural workers in small institutions. I propose to publish a few extra copies and begin by distributing the final finished version to the local cultural centers in my home town – the local archives and the aeronautical museum mentioned earlier in this paper. I can ask them if they know of any other local institutions that might like a copy as well and slowly spread my scope from there.

Another possible avenue I may pursue in order to distribute this booklet on a wider scale is to approach the Cultural Heritage Information Network, and see if this is a project they would like to work with me on. If this is a possibility I could expand the booklet to include a larger glossary, or more images.

There is also the possibility that I could create a series of booklets that progressively build on the topic. The next booklet could be the topic I originally started with – “So, you have a database...now what?” – in which institutions with existing collection management systems can learn about ways in which to get more use out of their systems and/or make their system more efficient. There could also be a booklet about how to ready your electronic records for on-line content. By this I do not mean how to digitize your collection (there is an abundance of these), but rather identifying copyright concerns, how to anticipate for a drop/rise in request for information/images, embedding metadata into images, etc. I feel that this project has endless uses and room for future growth due to its real world application and need.

Bibliography

Arneson, Arne J. & David Timm, "Content That Counts: How a University Library and a Public Library Digitized Local Content for the Web", Feb 2004.

[<www.scls.lib.wi.us/mcm/programs/digitize.ppt>](http://www.scls.lib.wi.us/mcm/programs/digitize.ppt) .

Besek , June M. "Copyright Issues Relevant to the Creation of a Digital Archive: A Preliminary Assessment", Council on Library and Information Resources, Jan 2003.

[<http://www.clir.org/pubs/reports/pub112/contents.html>](http://www.clir.org/pubs/reports/pub112/contents.html) .

Bullock, Alison. "Preservation of Digital Information: Issues and Current Status.", Information and Technology Services, National Library of Canada, April 1999.

[<http://www.collectionscanada.ca/9/1/p1-259-e.html>](http://www.collectionscanada.ca/9/1/p1-259-e.html) .

Canadian Heritage Information Network. "Collections Management Software Review – Comparative Analysis." 4th Edition, Canadian Heritage Information Network, 2003.

[<http://www.chin.gc.ca/English/Publications/software_review.html>](http://www.chin.gc.ca/English/Publications/software_review.html) .

Canadian Heritage Information Network. "Information Technology in Canadian Museums: A Survey by the Canadian Heritage Information Network.", Canadian Heritage Information Network, 2001.

[<http://www.chin.gc.ca/English/Reference_Library/Information_Technology/pdf.html>](http://www.chin.gc.ca/English/Reference_Library/Information_Technology/pdf.html) .

Green, David. "NINCH Guide to Good Practice in the Digital Representation and Management of Cultural Heritage Materials.", National Initiative for a Networked

Cultural Heritage, University of Glasgow, 2003.

<http://www.nyu.edu/its/humanities/ninchguide.pdf>.

Greenstein, Daniel, Bill Ivey, Anne R. Kenney and Bill Lavoie. "Access in the Future Tense", Council on Library and Information Resources, Washington D.C., April 2004.

www.clir.org/PUBS/reports/pub126/pub126.pdf.

Holm, Stuart. "Protecting Your Records: The Value of Collections Information.", Facts and Artefacts, 2nd edition, Chapter 13 *Data Security*. Cambridge: Museum Documentation Association, 1998. <http://www.mda.org.uk/protect.htm>.

Keene, Suzanne. "Digital collections : museums and the information age." , Butterworth-Heinemann, Oxford, 1998.

Kennedy, Randy. "Amassing a Treasury of Photography." New York Times July 20, 2005.

NISO Framework Advisory Group. " A Framework of Guidance for Building Good Digital Collections." 2nd Edition, Bethesda, MD National Information Standards Organization, 2004. March 1, 2006 <http://www.niso.org/framework2.html>.

Peters, Dale. "Digitizing Heritage", from Blurring the Boundaries / Training for Transformation , 27th February – 1st March 2002 / Iziko Museums of Cape Town <http://www.museums.org.za/mtconference/papers/html/peters.htm>.

Reinhard, Andrew. "Universals of Collection Management System Implementation", Willoughby Press, 2002. <<http://www.willo.com/services/articles.asp>>.

Santelli, David, Director. "Ensuring the Sustainability of Online Cultural and Heritage Content: From an Economic Model to an Adapted Strategy", Study conducted by The Multimedia World Watch at the request of Canadian Heritage 2003.
<http://www.pch.gc.ca/progs/pcce-ccop/reana/pubs/economic_model/economic_model_e.pdf>.

Simberloff, Daniel, Chair. "Long Lived Digital Data Collections: Enabling Research and Education in the 21st Century", National Science Board, September 2005. <<http://www.nsf.gov/pubs/2005/nsb0540>>.

Smith, Abby, ed. "Building and Sustaining Digital Collections: Models for Libraries and Museums", Council on Library and Information Resources, August 2001.
<<http://www.clir.org/PUBS/reports/pub100/pub100.pdf>>.

Waibel, Günter, "Along the Road to Shared Collections Access: on Integrating Collection Management Systems and Digital Asset Management Systems.", SPECTRA vol 29, issue 1, Spring 2002, pg 16-19.

Additional Resources

Reviews:

CHIN – Collections Management Software Review
http://www.chin.gc.ca/English/Collections_Management/Software_Reviews/index.html

- An excellent report and review of the different collection management systems available. Provides costs, and type of support vendor reliability, etc. If you are considering acquiring a system, or are simply interested in what else is available, be sure to check here first.

Standards:

http://www.chin.gc.ca/English/Standards/metadata_intro.html
- A wealth of information relating to metadata standards including Collection Description Standards, Data Models for Museum Information and an additional reading and resource section.

http://www.uniform.org/publications/uifm/may96_standards.html
Creating and Implementing Database Standards by Carl Cargill.
- From 1996, but it outlines why standardization within database technology industry is important and why users need to articulate their needs.

Digitization Guides:

Introduction to Imaging by Howard Besser, Getty Research Institute, 2003.

- Excellent guide and includes glossary. Available for free download
http://www.getty.edu/research/conducting_research/standards/introimages/

CHIN site under Creating and Managing Digital Content there are excellent in-depth resources available, such as cost, legal issues and standards. <http://www.chin.gc.ca>

Technical Guides:

Database Demystified by Andrew J. Oppell, McGraw-Hill/Osborne, Emeryville, CA, 2004

How Software Works, by Ron White, Ziff Davis Press, Emeryville, CA, 1995.

So, You Want A Database?:

A Beginner's
Guide to Database
Technology and
Collections Management
Systems for
Small Cultural
Institutions

Table of Contents

Introduction..... 3

What is a Database? 4

Why use a Collections Management System?.....9

Frequently Asked Questions 11

Glossary 18

Additional Resources 24

Acknowledgments: I would like to thank my thesis supervisor, Dr. Alison Nordström, for her wonderful advice and guidance, as well the numerous people who generously gave their time and knowledge.

Produced By
Shannon Perry
Ryerson Graduate
Thesis Project
Photographic Preservation
Collections Management
© Shannon Perry 2006

Query Language: Computer languages used to make queries or searches into databases. The languages will differ slightly depending on the system.

Scalability: A scalable system is one whose size can be adjusted to meet ongoing requirements or where each part of a system is flexible enough to accommodate growth or reduction in another part.

Schema: In computer science, a schema is a model and provides a means for defining the structure, content and to some extent, the semantics of documents. It is also a description of the structure of a database or a directory or a defined part of a database.

Usability: The quality of an application system that enables it to be easily understood and conveniently employed by the user.

network database has led to their wide use in business.

Normalized Data: Also called normalization, it is often used to create a set of tables from a set of data files and records and ensures that no fields contain duplicate information. This is necessary in order to ensure that a piece of information can be updated and the update "propagated" to the entire database without missing and instances.

Object Based Database: Also known as object-orientated database or programming. It is based on how we interact with objects in our everyday life. Take for example an apple, which has properties (size and colour) and capabilities (edible). In the other, older types of databases, the properties and capabilities were often grouped together. In the object based program, you are able to search for the object(s) using any one of the fields to describe its properties, and ... they store and manipulate more complex data structures called "objects" which are organized into hierarchical classes that may inherit properties from classes higher in the chain; it is the most flexible and adaptable.

Off-the-Shelf: Hardware and software that can be purchased and put in service without additional development costs or a type of no developmental software that is supplied by commercial sources.

Relational Database: A collection of data that is organized into tables so that relationships between and among data can be established. The relational model organizes the data as sets of tables where the column represents fields and the rows represent individual entities. A field (column) that two tables have in common can be used to link the two. The relational database uses the common field to join the two tables. It is used where associations among files or records cannot be expressed by links

Introduction

In the museum world, the word "database" is often thrown around in conversation. Talking about an institution's database often becomes a means of sizing up where the institution sits in the larger museum community. Questions such as "What kind of database and collection management system are you guys using?" or "Did you go with the additional interface that they were offering?" or "What is your record capacity?" can often make the conversation sound more like a car show than a museum- related conversation.

If your museum does not currently have an electronic database, or if you have only a limited understanding of what it does, these types of conversations can be overwhelming. Do we really understand what is being discussed, or are most of us simply repeating what we have heard from other staff members or the collection management company representatives? Perhaps if we take the time to understand what exactly a database system is and how to manage it, we can understand better what exactly it can help us do.

This booklet is not going to give a technical explanation, nor is it going to make someone into the resident IT expert overnight. The goal of this booklet is to help demystify the database and its surrounding bits and pieces and the collection management system, and in doing so, perhaps inspire you to pursue further study the subject and look towards using your database in more innovative and useful ways.

This booklet will cover such topics as budget concerns, how the database fits into the larger museum community and what you should consider for the future. It also contains a glossary of terms and a page of additional resources. Good Luck and Enjoy!

What is a Database?

So...what is a database? The term *database* refers to the organized collection of records. There could be as few records as three and it could still be called a database. The term database originated within the computer industry and technically, can include non-electronic databases within its definition - a card catalogue system would still be considered a database. Most people, when referring to a database however, are referring to an electronic database.

A database structure is fairly simple in concept. As illustrated in Figure 1, a character is the smallest element of storable data. A field is a group of related characters. A record is a group of related fields and a file is a group of related records, and a database – is a group of related files and records. If you want to be able to search your

database, or organize the records in any way, then you are actually talking about the Database Management system, or in the case of cultural institutions, the Collection Management System. The collection management system is software used to manage the database - whether it is an off-the-shelf product from a commercial supplier or an in-house created system. People often mistakenly refer to the entire management

are organized in a tree-like fashion, with each level of records branching off into a set of smaller categories. They provide single links between sets of records at different levels.

Metadata: Usually defined as "structured data about data". Metadata can include information associated with either the object itself, or the content of the object, for purposes of description, administration, preservation, the documentation of legal requirements, use and usage, etc. It can be stored internally – in the same file as the data, or externally – as a separate file. When museums are considering putting parts of their collection online, they will often embed the metadata into the digital object, to both provide information and prevent the object from being used an unwanted context.

Metadata Schema: A set of rules for recording or encoding information that supports specific metadata requirements or standards.

Migration: Digital preservation strategy that involves transferring data from a format or standard that is in danger of becoming obsolete to a current format or standard. The most common example is the process of upgrading files to become compatible with a new version of software or operating system. May also be known as conversion or reformatting.

Network: The physical hardware and software connections between computers that allow information to be shared and electronic communication to take place. A network printer, for example, allows many PCs to print to it, even though the PCs are not hardwired directly to the printer.

Network Model: Network databases create multiple linkages between sets by placing links or pointers to one set of records in another. The speed and versatility of the

Digital: Electronic technology that generates, stores, and transmits data - most commonly, as binary data (represented by a combination of 1's and 0's). Because there are only two possible values, the accuracy of binary digital data at any given point is relatively easy to test, helping to create accurate copies.

Digital Image: An image described as a set of digital data. Digital images may be digitized from traditional sources such as paintings or may be created digitally (such as an original digital photograph).

Digital Object: Considered the data and the metadata together as a single unit. It may also be known as a digital asset, an information object, or an information package.

Digital Preservation: The process of ensuring that a digital object is accessible in the future. It refers to the management of the digital object, which will require regular attention to ensure that the rapid technological advance do not leave you digital objects inaccessible due to out of date equipment or formats. This is still a new area, and many preservation experts are still unsure of how to handle the issue of digital preservation.

Digitizing: The process of deriving digital objects from their originals by converting them to binary code. Also known as image capture.

Documentation: Textual information that describes a work of art or image, recording its physical characteristics and placing it in context. It is also the printed and displayed materials that explain how to use or navigate the collection management system, including step-by-step written or pictorial instruction and standards. It may be regarded as one of the most basic preservation strategies for digital files.

Hierarchical Model: The records in hierarchical databases

system as the database, but it should be referred to as the collection management system.

The hope for a database and the collection management system is that the collection of records can be translated into pieces of knowledge. In order to make use of the various pieces of knowledge from the database (cross referencing objects, for example), there is a type of computer code known as a schema that is required.

The schema describes the objects that are represented in the database, and the relationships among them. There are a number of different ways of organizing the schema, that is, of modeling the database structure; these are known as database models (or data models).

The model in most common use today is the relational model which represents all information in the form of multiple related tables, each consisting of rows and columns (the true definition uses mathematical terminology which will not be discussed in this booklet). The relational model represents relationships by the use of values common to more than one table.

In cultural institutions, the model most often used is the object-based model, which is also known as object-orientated database or programming. It is based on how we interact with objects in our everyday life. Two people can describe the same object in completely different ways, and with the object based model, You are able to search for the object(s) using any number of fields used to describe the objects properties. It is based on the hierarchical model in that certain fields can have multiple sub sections/fields that allow for highly detailed searches.

Figure 3:
An example of
a search in an
object-based
database

Other models such as the hierarchical model use a more explicit representation of relationships. The records in hierarchical databases are organized in a tree-like fashion, with each level of records branching off into a set

can only be accessed and altered by those authorized to do so. It can also include the preservation of the data and assurance that the data remains in its intended form during and after retrieval, transfer, etc.

Data accessibility – Exactly as it sounds, data accessibility ensures that the software and data remains accessible to those who need it by ensuring that the programs remain up-to-date and that the various programs used within your institution remain compatible.

System development – The overall development and maintenance of the system. An average plan for system development can include a Feasibility Study, Analysis, Design, Implementation, Testing, Evaluation and Maintenance.

Database Model: The types of databases available, including relational, hierarchical, network and object based. There are three levels of database models: the general concepts and information level in which the database takes shape and information is grouped together in theoretical areas or levels. It is at this stage that the relationship between records and files is established. Next comes the Structured Data stage in which the facts and descriptions are firmly established. It is at this stage that the schema is created. Finally comes the actual stage of representing the data into structured code and creation of fields that have established relationships with other fields, etc., resulting in the model for the finished product.

Data Conversion: Data conversion is needed if you are switching from one system to another, or sometimes even upgrading to a newer version of the same system. Codes are written by computer programmers that essentially instruct the new program where the information from the previous program should go.

Glossary of Terms

Collection Management System: Consists of a database and facilities for adding, updating, retrieving, manipulating and reporting on data which document a collection.

Data: Although often associated with computers, data can actually be any type of recorded record. With computers, data is represented by a series of bits — a series of 1's and 0's— which are stored and saved in files within the hard drive. It is the numbers, text, graphics, images, and voice stored in a form that can be used by a computer. Data stored digitally is not a fact, but rather a representation of facts, and therefore accuracy and usefulness of the data depends not only on the accuracy of the original data but also on the context in which the data is being used within your database.

Database: Any structured collection of data or information that is specially organized for rapid search and retrieval. It can be organized into records with component fields and it can be digitized (in a computer) or not. A basic card catalogue system could be considered a type of database.

Database Administration: As it sounds, it is the management of the database system. It can be divided into four broad areas: data security, data integrity, data accessibility, and system development.

Data Security - is the means of ensuring privacy and making sure that data is kept safe from corruption and that access to it is suitably controlled.

Data Integrity – concerns the accuracy, correctness and validity of the data. It is the assurance that data

of smaller categories. They provide single links between sets of records at different levels.

Many IT professionals would consider a collection of data to constitute a database only if it has certain properties. For example - if the data is managed to ensure its integrity and quality, if it allows shared access by a community of users, if it has a schema, or if it supports a query language. In cultural institutions, there is also a strong emphasis on the inclusion of metadata, which is defined as "structured data about data".

Metadata can include information associated with either the object itself, or the content of the object, for purposes of description, administration, preservation, the documentation of legal requirements, use and usage, etc. It can be stored internally – in the same file as the data, or externally – as a separate file.

At this time there is no universally agreed definition of these properties mentioned above, although many museum communities around the world have created and strive to maintain a certain level of professional standard. Located in the Additional Resource section are websites which provide guides to museum data and metadata standards.

the system safe and secure in another location. It is important to make sure that the information within the server is updated regularly. If it is not possible to have a separate dedicated server, a small institutions can regularly save the information to a series of disks or other recordable medium and store the saved information in another location – in case there is ever an accident at the institution such as theft of the computers, fire or floods.

insurance information, an objects provenance and copyright material. What you *can* do, which will get more use out of your system is have the creator of your website pull certain pieces of information out of the database, such as the digitized images and accompanying metadata, and use it for the on-line content. You can update the website regularly as you add to your database, usually without any problem.

How do I regulate how/ when/ if data is entered into the database?

Most collection management systems will have some kind of security feature (user profiles) that will allow you to limit who is able to enter data or access certain parts of the system. Ideally there will be a database administrator. This person will be the only one who can perform global data changes, customize the system and work with creating, changing, and deleting user profiles.

Another important part of regulating how and what data is entered into the database is ensuring that the staff using the system is properly trained, especially in such areas as standardization. The reason this is important is that data standardization is key to making your system work, and the easiest way to keep the data standardized is to limit who enters the data and how.

How do I protect the information entered into the database in case of emergencies?

Every institution should invest in a dedicated server, which is a powerful computer (usually kept offsite) that is dedicated to running your collections management software, to storing your data, to keeping your collection's multimedia files in addition to the computers that are used in the day-to-day management of the database. This is important in case something happens to the computers in your institution or office, so that you will have a back-up of

Why Use a Collection Management System?

If you cannot easily locate the data for your collection within your database, the data is worse than useless, because it has a legitimate use that is not being met. If used correctly, collection management systems will help you access the information in your database and put it to use.

Collection management systems can often be expensive to buy, expensive to update, and hard to navigate without proper training. Depending on the software company, only a certain number of seats (actual terminals on which the collection management system can be accessed) may be available for use within the museum.

Preservation experts warn us about the need to stay abreast with the technology or risk losing years of work, and updating your technology every few years is a hassle many people would prefer to avoid. For all of the negative aspects, why would anyone want to use a collection management system?

A collection management system can store and access large numbers of records and information quickly and efficiently. A collection management system will allow you to cross reference objects or search using multiple fields, enabling you to perhaps identify relationships between objects not previously identified.

You will have greater room for information because the information stored within the computer hard drive is written in a compression code which allows you to store massive amounts of information in a minimum amount of space. Within the stored files you can easily add, edit or delete information.

A collection management system lets you arrange data to be pulled up according to certain specifications (i.e. 'find all Da Vinci items'). Security may be better than traditional paper based methods of cataloguing because you are able to create password protected areas.

By using a collection management system, you can reach most areas of the museum by allowing employees more access to relevant information. In doing so, there is the potential to use data within other museum related applications. For example, data can be used in checklists for upcoming exhibitions, or used as content for on-line collection. You can keep track of multiple faucets of an object in your collection, such as recording insurance information and values, copyright information and loans. You can attach similar or relevant files to an object record, such as justification reports and scans of archival news clippings, which would otherwise be separated from the object and most likely not used.

Bottom line is, with a collection management system you will get much more use out of your database and collection than before. It is important to understand however, that simply acquiring a collection management system is not going to make your collection more accessible or useful. If you are taking the time to set up a decent collection management system, ensure that the records and data and metadata are entered correctly and consistently.

On the other side of the argument, some institutions feel that employees working within the collection management system should know only the fields in which they are working daily, to prevent people from having access to, or changing records that they shouldn't – for example, insurance values. One way to solve this is to install security features, such as password protection to the areas where you do not want people changing records. It may require professional IT help to install, and may add more than a few dollars to your budget, but the benefit of a more secure system can be worth adding the additional features.

What does digitizing have to do with databases?

Everyone is aware of the enormous benefits that can come from having part or all of an institution's collection on-line or made available to the public electronically. After the institution acquires a collection management system, thoughts often turn to 'digitizing the collection'. Many people seem to feel that if you are entering object records into the database, it makes sense to create a digital copy at the same time. This is a good idea, but you will want to think out the project thoroughly and consider the extra costs and time required. There are several guides to digitization available on-line, which outline industry standards, and several of these should be consulted before any concrete plans are made. Located in the Additional Resource section are websites which provide guidelines to digitization and are worth looking at.

What information should I put on-line?

Many institutions that have websites also include a searchable database. Although the information looks similar to that in their database, it is not the actual database that is available on-line. There are areas of the database which you will not make public, such as

detailed written instructions on how to operate and navigate a system) and failure (not even knowing the passwords to gain access to the system). Also make sure that as many people as possible know how to work with the program – the more everyone knows how to use the collection management system, the more they will use it and remember how to use it if the documentation is ever lost.

Benefits include a completely custom system, and often, brilliant innovations come out of experimenting with, adapting or otherwise challenging the available collection management systems available at the time.

Does everyone on staff need to know how to use the collection management system?

Even if your museum purchases a simple or small collections management software tool, it will be investing money in implementation, training, etc. If the staff is sufficiently trained in the collection management system, then the program should help *any* museum professional work smarter and more efficiently while not demanding any changes to museum procedure.

In reality, what happens all too often is that a few people learn how to work in and/or navigate the collection management system, and the other staff members learn only the parts of the system that they feel may be relevant to their job or department. If they need to find something that reaches outside of this scope, often they will simply ask a more knowledgeable staff member to retrieve the information, rather than trying to learn for themselves.

One possible way to counteract this is to offer staff tutorials on how to work different components of the system at fairly regular intervals. Most people will become overwhelmed if told the need to learn how to operate an entire system in one lesson. By breaking up the lessons you can give staff a chance to become used to what they have learned and build on the existing knowledge with each successive lesson.

FAQ

In this section, I have tried to identify what the most frequently asked questions are regarding collection management systems.

Does every museum need an electronic database?

Ideally, every cultural institution *should* have some sort of electronic database and collection management system, because it is important to maintain records of all the things that add to the history and life of the objects in the collection (i.e. acquisition reports, loans, exhibits, etc).

Even the smallest institutions can benefit from having some sort of electronic database, because an electronic database and collection management system is usually the first step towards putting the collection on-line, where the potential for use is beyond measure. But first, you need to decide what the best system for your institution is.

What do I need to consider when devising a budget?

A big factor when considering what type of collection management system your institution should get is the cost. The cost will depend on whether you decide to go with an off-the-shelf model, an in-house system or a combination. There are so many variables when considering the cost of the system, that it is a good idea to look at the areas most people overlook when thinking about acquiring a new system.

Included in your budget – in addition to the system itself - should be plans for upgrades of both software and computers, training of staff, buying additional licenses if the price of the system you chose is dependent on seats, paying for annual support, and data conversion - if you are switching collection management systems (even from within the same companies if the version you were

previously using is outdated enough - probably a shorter time period then most people think).

Implementation costs do not stop once the program has been installed. If possible, save some extra funds in the budget for memory and operating system upgrades. In the age of ever increasing budget cut backs, it may be unrealistic to budget for the inevitable changes in technology, as mentioned above, but you should at least make the staff at your institution aware of the possible costs that may be incurred overtime.

Remember - the time and money spent working within the collection management system should at no point overshadow that spent on the collection itself. If this is happening, perhaps it's time to rethink your choice.

How do I decide what type of collection management system to use?

When deciding what type of collection management system your institution will acquire, you will need to re-examine the purpose of your collection, what you have in your collection and how you foresee using the collection – both in the present and in the future.

How many people will be using the system? You will need to speak with all of the potential staff involved to identify the most important features you would like to have in your system. You will need to realistically decide what your budget is – and stick to it.

If possible, talk with other institutions that have a collection similar to yours and see what system they are using. Ask to "try out" a working system in another institution to see if you like the feel of how the system works. You will need to use the collection management system everyday, so plan accordingly.

It would be a gross understatement to say that not all systems are created equally. There is an excellent appraisal report completed by the Canadian Heritage Information Network (CHIN) of the different systems

available which can be consulted when your institution feels it is ready to acquire or change an existing database and collection management system.

What is the difference between an off-the-shelf system and an in-house system?

"Off-the shelf" (a.k.a. "out-of-the-box") systems are often created by companies that work exclusively with museums and other cultural institutions, such as The Museum System (TMS) or Multi MIMSY. Personal systems would include systems such as Microsoft access and perhaps even FileMaker Pro. They usually offer some type of system support, and upgrades are often included for the first couple of years. There may also be some type of warranty, which is invaluable if something goes wrong with your system.

Possible downsides could be that the company supplied support can be inconsistent and missing an upgrade could leave you outmoded within a few years. A solution to these downsides would be to stay on top of the upgrades and establish a working relationship with the support staff if possible.

In-house systems are exactly what they sound like. They are systems that have been created by the institutions that will be using them, and are entirely tailored to the institution's particular needs and requirements. These are usually found only within larger institutions because unless you happen to have a computer programmer, software designer and or data modeler on staff, hiring someone to design and write a new system is costly. Downside to these systems - other then cost - can be that the programmers can quit, retire or die, leaving dismal documentation and employees who have no idea how to use the program/software.

One way to avoid these possible problems would be to make sure that all functions relating to the collection management system are fully recorded. Documentation can mean the difference between success (following