

Disseminating Water Information via CD-ROM: A Case Study

Terry Dodge, *Center for Environmental Studies, West Palm Beach, Florida, USA*

Introduction

The purpose of this paper is to discuss the development of an informational CD-ROM on water resources related websites and international water documents. This paper discusses the production of the CD-ROM, choices made, problems encountered and lessons learned. While there are many technical details, the paper focuses on why certain choices were made, products selected, and the philosophy and reasons behind creating the CD-ROM. The paper is intended to answer both the "how to" and the "why" questions behind distributing information on CD-ROM.

History and Background

In October 1998 a group of water and environmental information professionals participated in a workshop to discuss how to increase awareness of water issues, share information and better utilize resources. Twenty-five representatives from nine countries formed the core group of the WaterWeb Consortium (see <http://www.WaterWeb.org> and *Recommendations Report*, this issue). One of the main points made at this Workshop, especially from representatives of Latin American countries, was the lack of robust Internet connections and the need to create alternative methods of distributing information. That felt need, coupled with the desire to supply focused water related information, generated the idea of creating a CD-ROM containing information on water websites and other water related documents. The initial CD-ROM was intended as an experiment to determine the value of the information distributed and the practicality of using CD-ROMs as a distribution media. In January 1999, with a deadline of mid-March for distribution at the Inter-American Dialogue on Water Management (Dialogue III), we began the process of creating the WaterWeb CD-ROM.

Producing a CD-ROM for international distribution presented a number of issues and problems. Variables such as cost, presentation method, platform (PC, MAC, Unix), replication company, data acquisition method, disk organization, database design, art work, translations, and documentation had to be addressed. As preliminary research determined that cost would not be a major factor (a quan-

tity of 500 would cost less than \$2.40 each), the two most pressing issues facing the project were database design (so data acquisition could begin) and method of presentation. Given the distribution deadline, the database design question was critical as data acquisition could not begin without it.

Production of the WaterWeb CD

Two types of information were to be included on the CD-ROM: a database of water related websites (including descriptions, keywords, and key site features) and a series of international documents. The presentation on the CD-ROM had to be sophisticated enough to allow searching and easy review of search results and smooth enough for users to easily enter and exit documents. Accessing the data had to be simple and straight forward given that some of the audience for the product may not be computer literate. We had to have a fully searchable database but we could not require the user to buy or learn any sophisticated software programs. To open the information up to the largest audience possible, the machine specifications had to be kept to a minimum and smooth access to large documents and technical papers was a priority. As we were working under a very tight timeframe, any development of custom software was out of the question.

Choosing the presentation method was also critical. The decision of delivery program was uncovered when our CD-ROM replication vendor suggested Adobe Acrobat (see <http://www.adobe.com/supportservice/custsupport/download.html> for further information on products mentioned in this article). Review of another CD-ROM that was produced using Acrobat and the Acrobat file (.PDF) creation tools (Adobe Exchange, Catalog and Distiller) convinced us this was the appropriate method to use. Why was Acrobat chosen as the presentation method and not other options such as .HTML, word processing or other database software? The following factors influenced our decision:

- *Ease of distribution/installation* - Acrobat Reader is already a popular browser on the World Wide Web and both plug-in and installable versions of Acrobat

could be included on the CD-ROM (see the Adobe website above; the Reader is distributed at no cost and Adobe promotes its distribution). This would insure that everybody receiving the CD disk potentially would use the same version of the software. Additionally, it is available in multiple languages (Spanish and English in our case) and installation was easy.

- *Common, easy to use tool* - Many people are familiar with Acrobat Reader so there would be a base of users that would already be comfortable using it and for those that are not, the user interface is straight forward to learn.
- *Document formatting and ease of creation* - One of the purposes of the Adobe Acrobat software is to present documents with their original formats preserved on different platforms without requiring any use of the software program that originally created it (i.e. present the document just as it would originally look when printed). An Acrobat file looks just like the printed page. Creating .PDF (Adobe file format) documents is as simple as printing. Once the driver is installed it appears as another printer selection in the print dialog box. This means a .PDF file can be created from any software application that will print.
- *Search tools* - The Acrobat Reader has two search functions: one searches the document being viewed and the second searches an index of multiple .PDF files. A "hit" list of files containing the search string (keywords, phrases, Boolean, etc.) is returned. The Adobe Acrobat Catalog tool creates these indexes. It was clear that the Acrobat products had the search capability necessary for the project.
- *Document maintenance capabilities* - Using Acrobat Exchange software, .PDF files can be edited and links inserted to other portions of a document, other files, web sites and program functions. These features were necessary to create a smooth interface from the user documentation to both the links database and the other documents. The ability to establish a link to the execute menu items was especially useful. This allowed a search "button" to be defined, that when clicked on displayed the search dialog box.
- *Portability* - Adobe Acrobat is meant to provide portability of information between systems. Moving information from a network hard drive to a CD-ROM to another users' CD-ROM is not dependent on a specific drive letter designation. If the directory structure remains intact, the data is accessible and the links work. Versions of Acrobat reader are available for all major platforms (Windows, MAC and Unix), so if a CD-ROM is formatted to be readable for both PC and MAC platforms, the data should be accessible to all users.

While the database design issue was critical, we had experience with these issues as the Florida Center for Environmental Studies (see <http://www.ces.fau.edu> for

more information) had already designed a Links database (on-line at <http://www.waterweb.org/linksdb/>). In designing that database, the Dublin Core metadata standards had been adapted (see <http://purl.oclc.org/dc/> for more information). Our basic design had already proven itself to be a versatile and flexible format. We added some specific water related categories (for ease of searching) and website features to allow the database to meet the specific needs of the intended audience.

The Links database was developed in Oracle with an .HTML interface. This allowed records to be added and edited via the Internet (on-line). This interface was critical as we rely on volunteers from widely located organizations to contribute content (descriptions of websites). In fact, data acquisition turned out to be the hardest part of the entire process, not because of technical difficulties but because of volunteers scarcity of time to devote to the project and the projects very tight time schedule. However, after two weeks a minimum number of entries had been received.

Our time frame was partly dependent upon the company chosen to replicate the CD-ROM. As it turned out, local resources (West Palm Beach, Florida) did not produce a viable vendor, but a web search uncovered several possible vendors. After contacting several, reviewing their websites and speaking with references, DiscPress of Waterbury Connecticut was selected. Their typical turn around time was approximately two weeks (without any premium payment). They accepted art work and data via electronic submission, they supplied packaging, and if necessary they were able to provide other services.

Design

Having made these major decisions it was time to turn attention to the details of designing the CD-ROM. Some were obvious considerations and some were afterthoughts, but all had to be addressed to successfully complete the project. Some of the technical computer issues that had to be determined were: which network structure to use for our work area (we setup a shared drive), disk directory structure and file naming conventions, inclusion of Acrobat program files, and the linking of documents, web sites, captured home pages and commands. The technical issues turned out to be straight forward and easily solvable.

The feature of including captured home pages on the CD-ROM database so that users without Internet connections could view the websites home page was decided upon early on. The biggest problem encountered in creating .PDF files of these home pages was that any background images or colors on it were not displayed in the file. This usually is not a problem, but for any websites using white text on a relatively dark background, the text may be lost. In general, we found the features of Adobe Acrobat provided the capability to solve our technical issues. (In this case, since some links on the CD-ROM were to docu-

ments to be contained on the CD-ROM and others were to websites (external), we designated the two different types with different colors, blue for external, red for internal).

The artwork (see Figure 1) and the CD-ROM packaging were relatively easy issues; we used a window sleeve package that let the CD-ROM art work show through thereby eliminating the need for a more expensive plastic case. The logo from the WaterWeb.Org website (from <http://www.waterweb.org>) was used along with some simple text directing users to view the Readme.txt file.

Execution of Project

The most difficult and time consuming production issues were: document acquisition (some documents were not available until the deadline and then had to be processed into .PDF files), creation of the introductory and instructional documentation, and translation. In particular, the last two items turned out to be very time consuming. We were able to find translators, but there was some reluctance to translate documents with technical terms and jargon, both technical water resources and computer terminology. We learned that if you must have a technical document translated, it will take additional time to find a highly qualified translator. The documentation included on the CD-ROM consisted of initial instructions printed on the CD-ROM, a text readme file, a startup file which included the table of contents, usage instructions and other reference material, a comment form, an alphabetical list of web sites (linked to database entry), a list of water resources listservers, and WaterWeb web ring information. The startup file demanded the most work. Unfortunately Acrobat Exchange is not meant to be a powerful text editing tool, so when the startup file needed significant

changes, they had to be done in a word processor and then a new .PDF file had to be created. This became tedious as it meant that the links within the startup file had to be redone (approximately 20 links; it wasn't an unmanageable task, but with the necessary testing procedures, it became a tedious task).

The contents of the startup file were very critical if use of the CD-ROM was to be smooth and easy for its users. The usage instructions had to be clear and concise to an international audience. The technical information had to be included for those interested users and it had to be logically organized. A screen capture tool was used to create pictures of portions of the Adobe Acrobat Reader screens to include in the documentation. The instructions detailed the keystrokes to perform certain functions but hot links were also used to activate these program functions. This allowed us to have a "click here" link that brought up the search dialog box. Consequently the user did not have to find the Reader icon or navigate through many menu choices. Providing both the needed information (both technical and topical) and easy access to that information were the hallmarks of our project.

Two other tasks proved to be time consuming and labor intensive: capturing the water resources website home pages and linking database entries to both these captured home pages and the active web site (the URL). At our cut-off date for data acquisition, we had 168 websites in the database; this meant someone had to visit the 168 sites, "print" the home page to a file, and then edit the individual database record to add two links: one for the actual web site and one to the captured home page. A student intern provided the labor for this part of the project. A linked alphabetical list of sites was also created. The most important parts of this task were to provide a naming convention for the captured files (the file name corresponded to the database identification number) and testing all links; if even one was not accurate, the integrity of all the data would be called into question.

Having created all of the pieces for the CD-ROM, only the final task of putting it all together remained. A common word processing package produced the .PDF files for the international and instructional documents. The Oracle database provided the individual .PDF database files and the alphabetical list. Adobe Exchange linked everything together and Adobe Catalog created the index of the database entries. All these files resided on a shared drive on our network. To actually write the CD-ROM, all that was necessary was to copy the entire drive to a CD-ROM writer. One helpful feature of the CD-ROM writer software is the capability of being able to update or add more information to the CD-ROM after its initial creation. This allowed us to fully test the CD-ROM, then correct or change files, and then write the changed version back to the original disk; this was especially useful with the startup file). When everything has been written to the CD-ROM as desired, the CD-ROM is "closed"

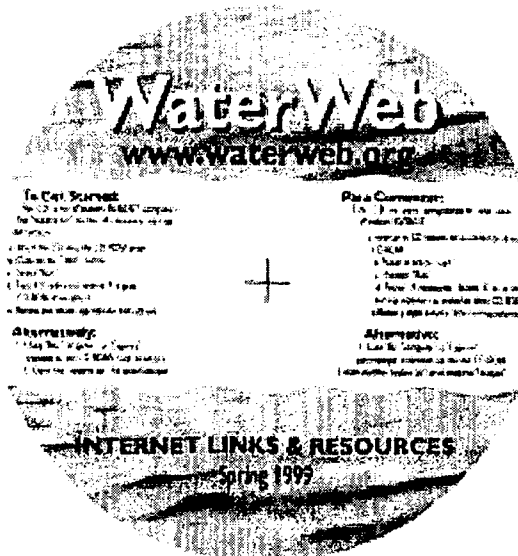


Figure 1. CD-ROM

blocking any further writing to it. The "closed" CD-ROM is then submitted for replication. The transfer films for the artwork were submitted with the CD-ROM to our vendor.

Lessons Learned from this Project

What problems did we encounter and what lessons did we learn during this project? As indicated earlier, data acquisition was the biggest and most time consuming issue; as with most projects that depend upon human cooperation, more time should have been allocated for its completion. While systems problem did take the on-line submission of data down for six days, the real problem was finding qualified water resources volunteers willing to review and document web sites for the database. The other major issue also had a human dimension: the translation of the documents, mostly into Spanish. Translation services may be easily obtained but for the best results it should not optimally be done under pressure or time constraints. If technical material, whether computer or water resources related, needs translation, significantly more time must be allowed to find qualified people and have other native speakers review the translation.

The technical problems encountered in this project were minimal and we found them to be easily dealt with. The issue of web site backgrounds not displaying on .PDF files was something we lived with and did not affect many web sites significantly. The limited ability to maintain .PDF files led us to streamline the startup file and we became proficient at defining links in that file. In a few cases, documents were broken into multiple .PDF files for ease of maintenance. Given the strong technical team we had, computer related problems were not experienced as a major barrier to implementing this project.

Even with the tremendous growth of the Internet and World Wide Web connectivity, information distribution is still a problem, especially in regions with limited on-line access. Many areas of the world do not have robust communication systems or cannot afford unlimited Internet access. Anyone who has tried to find worthwhile on-line water related information agrees that it can be a time consuming and frustrating effort. The main value of using CD-ROM technology to distribute information is that they are very portable, they are relatively cheap and easy to produce and they can be tailored to meet the demands of the user. The value of the WaterWeb CD-ROM is that it provides very focused, high quality information, it points people to consistently valuable websites and its computer system demands are minimal. While the Internet gets all the attention in today's information economy, CD-ROM technology has a place as well and can serve users well when applied to the appropriate task.

About the Author

Terry Dodge is the Information Services Coordinator for the Florida Center for Environmental Studies (a water focused research center). After earning a Bachelors degree in Business Administration from the University of South Florida, he has spent 22 years in the computer industry with the last seven conducted in an academic environment. His current duties include coordinating all information services for CES, overseeing eight web sites and actively working with groups to improve information accessibility. Projects with the Everglades restoration and water management in the Americas currently take priority. The author can be reached at TDodge@centauri.ces.fau.edu.

Discussions open until December 31, 1999.