

Publishing a newspaper index on the World Wide Web using Microsoft Access 97

Maureen Kilcullen

Describes conversion of a database-created index into a suitable format for publication on the Internet. The index to the *Canton Repository* was converted from dBASE IV and uploaded in Access 97

Background

Kent State University Stark Campus librarians began indexing the local Canton, Ohio newspaper in 1993. The *Canton Repository*, originally issued under the title, the *Ohio Repository*, has been consistently published since 1815 and documents the history, social life and customs of northeastern Ohio. The librarians knew from the outset that the index would interest not only the university but also local academic, public and special libraries and local historical societies. Since the librarians wanted to make the information available in print and electronically, they decided to index the newspaper with the help of a computer. After reviewing a number of software programs, a DOS-based version of the software program dBASE IV was chosen because of its ability to construct a subject authority thesaurus as well as its ease of use. The structure of the database was designed using alphanumeric character fields: heading; subject; name; byline; abbreviated newspaper title, page, column; and the date. dBASE IV proved to be an excellent choice for that time and made the job of inputting and organizing the data much easier (Kilcullen 1996).

Implications of new technology

Although the plan was to provide the information in print or electronically, ideally, the librarians wanted to make the index available over the Internet.

Maureen Kilcullen, Reference librarian for Kent State University Stark Campus, had been working almost exclusively with the Internet and realized that with the right technology, the database could be offered as an Internet FTP (File Transfer Protocol) file (Gilster 1993). Soon though, innovations and technical breakthroughs in the World Wide Web and Microsoft Windows software began to affect this thinking — rather than offer the index as an FTP file, why not publish it on the World Wide Web? The popularity of the Internet continues to steadily increase partially because of technological advances in the World Wide Web (Walsh 1996). Microsoft software developers have kept up with these advances as has Kent State University.

New Dell computers operating on Windows 95 and loaded with the Microsoft Office Suite were purchased for the library staff in 1996. Microsoft Office Suite includes a number of programs that can create HTML (HyperText Markup Language) files to be used on the World Wide Web, specifically Microsoft Word, a word processing program, and Microsoft Access 97, a database management system (Lemay 1996).

Kilcullen had used Microsoft Word to create and edit HTML

files for the library and was intrigued by Access 97's abilities to import dBASE IV files and then export the data in HTML format to the World Wide Web.

One of Access 97's strong points is its ability to transform existing database table files created by DOS applications (such as dBASE IV) into the Microsoft Access 97 format - a process called importing a file. Access 97's better Internet integration lets users output tables, reports or datasheets to an HTML file.

In addition, Microsoft Access 97 also features useful tools called Wizards which "are special dialog boxes that ask you questions about the document you want to create and then use your answers to layout and format the document" (Using Microsoft Office c1994).

Converting the index

Kilcullen was concerned that importing the index from dBASE IV to Access 97 might corrupt the file, but after researching the two programs she found several similarities. For example, both dBASE IV and Access 97 are computer applications that can retrieve, sort and print information in a database. The data in both programs appears in a table format and both use similar terminology: column headings are called "field names", the columns are "fields" and the rows of data are "records" (Using Microsoft Office c1994). Both Access 97 and dBASE IV field databases are similar in that dBASE IV uses "alphanumeric characters" while Access uses "alphanumeric text" (Jennings c1994).

After this investigation, Kilcullen felt confident that the index would not be corrupted if imported into Access 97. She then began the process to import the index from dBASE IV into Access. It was very simple.

To import the index into Microsoft Access 97:

1. Open Microsoft Access
2. A dialog box appears and offers the following choices:

Do you want to open a Blank database?

Do you want to open a Database Wizard?

Do you want to open an existing database?

3. Choose "Open a Blank database" and a dialog box appears and asks for a filename, drive and directory in which to save the database. After these choices are made, click on "Create".

4. The next dialog box appears with tabs marked:
- tables
 - queries
 - forms
 - reports
 - macros
 - modules.

With the tab set on "tables", click "New".

5. A dialog box appears for the new table and offers the following choices:

Datasheet view: "the default method of displaying a table".

Design view: "displays the characteristics of each field on a grid format similar to a spreadsheet".

Table Wizard: "chooses the fields for your table from a variety of predefined tables such as business contacts, household inventory, or medical records."

Import Table: "Imports tables and objects from an external file into the current database."

Link Table: "Creates tables in the current database that are linked to tables in an external file".

(Instructions above based on Jennings c1994).

Highlight "Import table" and then Click "Ok".

6. Find the database file; highlight it and click on "Import".

7. Access 97 creates a new Access table in which the imported data is stored.

The import was successful and was as easy to edit and update as in dBASE IV. The keyboard function in Access 97 (ctrl + ' [apostrophe] or ctrl + " [quotes], inserts the value from the same field in the preceding record) is similar to the use of the function key f8 in dBASE IV.

A major advantage Microsoft Access 97 has over DOS-based dBASE IV is Access's ability to sort by field without creating a new database. In dBASE IV if one wanted to sort the index by date or name rather than subject, one would have to sort the file on that field name to another new database file. For example in dBASE IV, the command would be: "sort on date to date.dbf" which sorted the records in either ascending or descending order and created another database file named date.dbf. In late 1996, with over 5,000 records in the file, memory would shortly become a premium if this method continued to be used. Access 97 allows one to sort the records on the different field names (heading, subject, date, etc.) without creating a new database file. One simply clicks on the appropriate field header then clicks on "Records" in the toolbar, then clicks sort in ascending (A-Z) or descending (Z-A) order. The table is quickly sorted and does not create another new index unless it is saved. This option is very useful when editing the index.

Publishing to the Web

The time came to experiment with publishing a sample of the index on the web. Since Access 97 creates a web page for each page of a report, Kilcullen decided to keep the table format,

supported by both Netscape Gold and Microsoft Internet Explorer. The table format also lines up the records so they can be viewed easily on the web. Publishing the index on the web was made even easier by using Web Wizard dialog boxes. The steps to publish the index to the web are as follows:

1. Open the Access 97 database file to be published to the Web.
2. Click on "File" in the toolbar, then click on "Save as HTML"
3. The "Publish to the Web Wizard" dialog box appears with step by step instructions. Click "Next" to proceed.
4. The dialog box asks whether to publish a table, query or form. Select the appropriate format (in this case, table) and click "Next".
5. Choose whether or not to use an Access 97 template to publish the document. (You can use or modify provided templates or build your own — Kilcullen chose not to use a template.) Click "Next" in the next wizard.
5. Select "static HTML" if your data rarely changes after it is published or "dynamic HTML" if you want the data to reflect changes. (Kilcullen chose "static HTML") and click the "Next" button
6. Choose the drive and directory you want to publish to and click on "Next", then click "Finish". The Web Wizard publishes the document to an HTML format.

Publishing the sample of the index to the World Wide Web was a simple and successful endeavor with the use of Microsoft Access 97. Thus far, the only drawback has been the lack of a search engine to navigate the file. The Kent State University Stark Campus network runs on a Novell 4.11 web server and does not support CGI (Common Gateway Interface) or Perl (programming language) scripts (the normal programming that features search capabilities) and has not been able to offer a search feature (Lemay 1996).

Future plans include upgrading the web server to support a search function, but in the meantime, the index has to be browsed or searched by using the "Find" button in Netscape. Kilcullen plans to continue publishing the index to the World Wide Web. To view the sample, go to:

http://www.stark.kent.edu/library/new/bydate_1.htm

References

- Gilster, Paul (1993). *The Internet Navigator*. New York: John Wiley & Sons, Inc. 83.
- Jennings, Roger (c1994). *Using Access 2 for Windows*. Special ed. Carmel, Ind.: Que Corp. 136.
- Kilcullen, Maureen (1996). "Indexing a local newspaper using dBase IV." *The Indexer* 20 (April 1996). 17.
- Lemay, Laura (1996). *Teach Yourself Web Publishing With HTML 3.0 in a Week*. 2nd edition. Indianapolis, Illinois: Sams.net Publishing. 350.
- Using Microsoft Office (c1994), special edition. Indianapolis, IN: Jennings op cit. 68.
- Walsh, Aaron E (1996). *Java for Dummies*. Foster City, CA: IDG Books Worldwide, Inc., 12.

Maureen Kilcullen is a Reference Librarian for the Stark Campus of Kent State University.