A project to develop an automated index of information about existing computerized instruction in the health sciences is reported and described. Methods of obtaining and indexing materials for the catalog are detailed. Entry and recovery techniques and selection of descriptors are described. Results to date show that the data base contains information on 226 units of computerized teaching material, broken down into categories of Medicine, Dentistry, Nursing and Pharmacy. Plans for broadening the base and updating it regularly are discussed. (SK)
BACKGROUND

A project is under way, supported by the National Library of Medicine, to develop an automated index of information about existing computerized instruction in the health sciences. Part of the need for such a project arises from the fact that good CAI materials are expensive to develop, and most health sciences teaching institutions could appropriately use more CAI programs than they have the resources to create. If existing CAI materials could be shared between institutions, then the cost per institution would be lower and use of computerized methods in the curriculum would be facilitated.

There are a number of difficulties that impede the sharing of CAI materials. The most common problems of sharing are lack of faculty acceptance, language and system incompatibility, and lack of workable release policies. At a very basic level, the health professions educator has the problem of locating computerized teaching materials that already exist at other institutions and might be appropriate for his students. A catalog would be useful, and should include enough information about each teaching unit to allow a preliminary evaluation of its applicability to another specific setting.

Since this is a changing and expanding field, the catalog must be frequently updated if it is to reflect an accurate picture of this body of CAI materials. Revisions and updating can be done at any time and with little effort if the information in the catalog is stored in an on-line computer system. The computer can be programmed to produce a printed version of the catalog, and this can include several types of indexes to the contents.

PREVIOUS WORK

The idea of a catalog of CAI materials is not unusual, and is similar in many ways to catalogs or indexes of printed teaching materials. Several CAI catalogs have been developed and published, such as the one originated by Helen Lekan at University of Wisconsin-Milwaukee. In addition, most institutions that develop computerized teaching materials maintain some form of internal documentation of their own programs although this varies in quality and completeness.

The first collection of information specific to CAI in the health professions was called "Guide to Computer Assisted Instruction in the Health Sciences". This was done by Brigham, Kamp, and Cross and published in December 1972 with citations for 362 health sciences CAI programs. For each computerized teaching unit, this catalog contained the name of the unit, the name of the author, the programming language used, and the institution where it was developed. More information was
needed before an instructor could make an initial judgement about the suitability of a particular CAI program for his courses and students.

GOALS OF THE PROJECT

One goal of this project is to update the existing information about health sciences CAI materials and enter it into a machine-readable data base so it can be maintained and summarized more easily. The data is being entered into an on-line text editing system (IBM's Administrative Terminal System), and additions, deletions, and corrections are accomplished via a computer terminal.

Another goal is to collect more information about each of the computerized teaching units. The initial survey (Brigham et al. 1972) included only minimal data. The current project uses a "Course unit information form" which has categories for Subject, Institution, Intended Audience, Computer Language, Name of the Unit, Author, Type of Computer, Instructional Strategy, Completion Time, Type of Terminal Used, Other Materials Needed, and a brief description of the program.

For each category of information, a description of the desired type of data was written. For example, in the "Strategy" category the description reads: "How does the program interact with the student? Some of the strategies commonly used are drill and practice, self-evaluation, tutorial, problem-solving, and simulation." A sample of the Course Unit Information Form and the accompanying instructions for its use is included at the end of this paper.

The final goal of the project is to use this machine-readable data base of information to create a catalog with key word indexes for various categories of data. In a planning survey, the Subject Area of the teaching unit was overwhelmingly chosen as the most important category to be indexed. The information contained in this category is in the form of a short keyword title for the content of each course unit, and this is used to generate a keyword index. Thus a title like "Pathology of the Liver, Neoplastic Disease" would be indexed under "Pathology", "Liver", and "Neoplastic" and could be located with a variety of descriptors.

PROCEDURE

Data gathering for this project was done in two stages. First, the institutions known to be active in the development of computerized health sciences teaching materials were asked to supply more complete information about materials already cataloged, and to give new information about teaching units developed since the last survey was done. In addition, requests for information about CAI development were sent to all schools of medicine, nursing (degree schools only), pharmacy, and dentistry. The resulting data was entered into an on-line data system in a standard format corresponding to the Course Unit Information Form.

The data is stored in two sections, one of which contains information about the teaching units. One entry is made for each teaching unit, with an average length of 20 lines. A separate section of data was
established for information about institutions. This is organized with one entry per institution, and contains the data about the computer installation and the CAI contact person(s).

The organization of the project is shown in this diagram:

```
DATA ENTRY
VIA TERMINAL

EDITING AND
UPDATING VIA TERMINAL

ON-LINE
DATA BASE

CATALOG WITH
KEY-WORD INDEXES

FORMATTED
LISTING OF
ENTRIES

DATA EXTRACTION
AND
MANIPULATION

DISTRIBUTION
TAPES
```

Once the data base is established, a number of different operations can be carried out whenever desired. A key-word indexing program is used to create a catalog and one or more key-word indexes. This program, called QUIC, can create key-word indexes on any or all of the categories of information on each entry. Normally, indexes will be generated for Subject, Author, Computer Language, and Institution.

Utility programs are used to alphabetize and number the data base. Simple PL/I programs produce formatted listings of the data or extract and tally certain categories of information. Tapes can easily be made so the entire data base can be transported to another center. All of these operations are run as batch jobs, initiated via job controls submitted from the on-line editor.

RESULTS

At the time this paper is being presented (January 1975) the data base contains information on 226 units of computerized teaching material. These break down into categories of Medicine, Dentistry, Nursing, and Pharmacy, depending on the school where they were authored. Two additional categories are needed to account for the rest of the teaching units. The "General" category includes health sciences teaching materials that are suitable for all health sciences students, and the "Other" category includes units that are in other health sciences topics such as medical technology, veterinary medicine, or patient education. The number of units reported in total or in each category is a potentially misleading figure, since "teaching units" may vary from 10 minutes to 50 hours in length. For this reason, the number of hours of teaching material will also be mentioned when discussing the data.
Attempting to classify these teaching materials also brings out the fact that some of them may be suitable for audiences other than the one for which they were written. For example, some of the teaching materials written for medical students are also suitable for nursing students. To find the total number of teaching units for nursing students, one would have to add those units in the "Nursing" category, those in the "General" category, and those in the "Medicine" category that are also suitable for nursing. The following table shows the number of units and the number of hours of computerized teaching material in each category. Figures in parentheses indicate additional units or hours from other categories that are applicable.

<table>
<thead>
<tr>
<th>Category</th>
<th>Units</th>
<th>Hours</th>
</tr>
</thead>
<tbody>
<tr>
<td>Medicine</td>
<td>107</td>
<td>233</td>
</tr>
<tr>
<td>Dentistry</td>
<td>44</td>
<td>84</td>
</tr>
<tr>
<td>Nursing</td>
<td>18 (+7)</td>
<td>53 (+8)</td>
</tr>
<tr>
<td>Pharmacy</td>
<td>9 (+8)</td>
<td>27 (+53)</td>
</tr>
<tr>
<td>General</td>
<td>14</td>
<td>31</td>
</tr>
<tr>
<td>Other</td>
<td>34</td>
<td>151</td>
</tr>
</tbody>
</table>

Overall totals show 226 units and 579 hours of teaching materials. For medical students there are a total of 267 hours (233 + 31) of teaching materials, and for dental students there are 115 (84 + 31) hours. For nursing students there are 92 hours (53 + 8 + 31), and for pharmacy students 111 hours (27 + 53 + 31).

The largest group of computerized teaching materials has been developed for medical students, and in this group the most frequent topic is general medicine. The data base shows 36 units (36 hours) of teaching materials in general medicine, 14 units (63 hours) in pharmacology, 11 units (17 hours) in physiology, 7 units (21 hours) in cardiology, and 6 units (33 hours) in anatomy. Other topics fall into smaller groups.

The next largest group of teaching materials is in the field of dentistry, where the most frequent topic is clinical dentistry. There are 11 units (42 hours) in clinical dentistry, 7 units (16 hours) in histo-pathology, and 5 units (8 hours) in anatomy. The next most frequent topics are dental materials, pediatric dentistry, and business management.

An attempt was also made to judge the proportion of the teaching materials that could be called "clinical" in their orientation. If "clinical" is defined as being related to case studies, diagnosis, and patient management, then about 35% of the 233 hours of specifically medical materials falls into the clinical category. The corresponding figure is 75% for the dental materials, 55% for the nursing materials, 65% for the pharmacy materials, and about 10% for the materials suitable for all groups of students (general materials). Grouping
these teaching units into categories has been done with the best judgement of the author, and is subject to the potential errors inherent in judgements made by one person.

CONCLUSIONS

Based on the current (January 1975) information about existing computerized teaching materials in the health sciences, one immediately notes that 362 units of teaching materials were reported in the 1972 survey by Brigham and only 226 units have been reported in the current study. Part of the difference arises because the current data gathering is not quite complete. There are estimated to be 25-50 more computerized teaching units that have not yet been reported. Another reason for the decrease in number of units is that teaching materials that were reported as individual units (such as 33 units for the Pilot Medical School at Ohio State University) in 1972 are reported as a single unit in the current study.

To take a specific example, there were 18% units of instruction in medicine reported in the 1972 study, and the current data contains 121 units (107 Medicine plus 14 General). Of the original 18% units, 73 are still present. There were 74 units reported in the earlier study for which there is no information in the present data. This may have been caused by discontinuation of CAI projects or by some difficulty in the current data-gathering process. Most of the rest of the original 18% units have been condensed into a smaller number of units. There are currently 48 new units in medical topics that did not appear in the 1972 study. The only other category that is comparable between the two studies is the units of dental teaching materials, which increased from a reported 38 in 1972 to 58 in the current study.

Another conclusion that can be reached at this stage of the project is that gathering this type of data is more difficult than it might seem. In order to produce a useful catalog, the data on each unit of teaching material must include a significant amount of detailed information. The work of providing this information may require a lot of time from staff persons who are busy with other projects having higher priority for them. Since most institutions developing computerized teaching materials have some form of internal documentation, the task of data gathering would be eased if these existing descriptions could be used. Some of the institutional documentation is of very high quality, but there are always differences in emphasis and in the particular categories of information that are included. An agreement on a standard form of documentation for these units of computerized teaching material would greatly facilitate the development and maintenance of a data base such as the one described in this paper.

PLANS AND RECOMMENDATIONS

The collected information will be published in the form of a catalog with key-word indexes. Publishing will be done on a non-profit basis by the Health Sciences Interest Group of the Association for the Development of Computer-based Instructional Systems (ADCIS) and should be under way by April of 1975. Availability will be widely announced,
including the ADCIS newsletter, appropriate journals, and mailing lists associated with this project.

Based on the large number of requests for information about this project and for a copy of the resulting catalog, it seems likely that this type of current data will be useful to health sciences educators. Since the data is now in machine-readable form and can be easily maintained, the periodic maintenance and updating needed to publish a new edition of the catalog would not require a large amount of support. The project should be continued, preferably by a national agency such as the National Library of Medicine which would be in a good position to develop and encourage the use of a standard format for information.

This project is supported in part by a grant from the National Institutes of Health, National Library of Medicine #LM-01855.
GUIDE TO THE COURSE-UNIT INFORMATION FORM

Please use a separate form for each course unit you describe, and include only course units that have been developed or significantly modified at your institution. Do not include course units that have been developed elsewhere and are in use without significant modification.

SUBJECT: A short key-word title for the content of the course unit. This information will be used to generate a key-word index, so a title like "Pathology of the liver, neoplastic disease" would be indexed under "pathology", "liver", and "neoplastic disease". This allows a course unit to be located by looking it up under any one of several categories.

SOURCE: The name of the organization or institution where the course unit was developed, with the name of the school and/or department where possible. I.e. "Ohio State University, School of Nursing". This information will reference a separate collection about each institution, including the contact person and the computer installation.

AUDIENCE: The level of training and the type of student for whom the course unit was developed. For example, "Medical students, preclinical", or "Nursing students, post-graduate", or "Physicians, continuing education".

LANGUAGE: The computer language in which the course unit is written. It will be helpful to include the complete name of the language, such as "Coursewriter 3 version 3".

NAME: Many course units have a one-word code name, such as ACIBA, ORALCA, CASE4, MACPEE, or DCLSB2B. The name may not have a descriptive function, but helps to identify the course unit, so please include the name if there is one.

AUTHOR: The person or persons responsible for creating the course unit. If the course unit was originally created elsewhere but has been significantly modified at your institution, give the name(s) of the person(s) responsible for the version you are using.

COMPUTER: The make and model of the computer central processing unit for which the program was designed.

STRATEGY: How does the program interact with the student? Some of the strategies commonly used are drill and practice, self-evaluation, tutorial, problem solving, and simulation.

AVG. TIME: The amount of time taken by most students to go through the course unit. If there is no definite end point, give the amount of time most students spend with it.

TERMINAL TYPE: If the course unit is suitable for a variety of terminals, please give the type(s), such as typewriter, CRT, or graphic, and the speed(s). Otherwise give the make and model of the terminal that is used. Please mention other equipment if it is required, such as a light pen, slide projector, movie projector, or audio unit.

OTHER MATERIALS: List other materials needed by the student, such as manuals, slides, models, video cassettes, etc.

DESCRIPTION: A brief (under 100 words please) account of the content and how the course unit is used by the student. You may want to include other information not covered by the preceding headings, such as a history of its use, availability, whether instructional objectives are available, or whether evaluation has been done.