Creation of a Forest Science Data Bank

Susan G. Stafford, Paul B. Alaback, Greg J. Koerper, and Mark W. Klopsch

The Forest Science Data Bank (FSDB), which serves the Department of Forest Science at Oregon State University, has evolved over more than a decade. Faculty and students have collected sizable amounts of data from silvicultural and genetics studies conducted throughout the Northwest. In addition, research at the H.J. Andrews Experimental Forest, formerly supported by the U.S.–International Biome Program and currently a National Science Foundation Long-Term Ecological Research Site, has produced vast quantities of data. Before the data bank was established, however, most data sets had incomplete or nonexis-

tent documentation; furthermore, to retrieve data, the user needed prior knowledge of the data set.

Our basic goal has been to document enough information about older studies to ensure their future value and to establish procedures for acquiring new data sets with complete documentation. To do this, we centralized quantitative services and pooled resources from various research projects. We identified specific tasks associated with the various aspects of data management and statistical analysis and recruited personnel with corresponding skills. Our staff now includes a statistical consultant, a data bank manager, and two data analysts—all of whom have research experience—and four data-entry personnel. *Figure 1* depicts the conceptual structure of the FSDB and its relationship to research and to information retrieval.

To collect new data, we use project-specific field sheets that facilitate data processing and analysis without compromising the efficiency of field measurements (Wagner 1981). A continually updated master file stores and organizes information about each experimental site, allowing data to be accessed and retrieved by plot. To preserve the value of these stored data, we require documentation on both data structure and study methodology for each data set. Also, access to, and selection of, data sets must be convenient for users unfamiliar with the breadth of data available. Our system includes descriptive, written information on the research origin of each data set and lists key words useful for search and selection.

We store and access written, descriptive study documentation using DataStar™ (Micropro International,

NEW
FORESTS
FOR A
CHANGING
WORLD

Proceedings of the 1983 Convention of
The Society of American Foresters,
Portland, Oregon   October 16-20
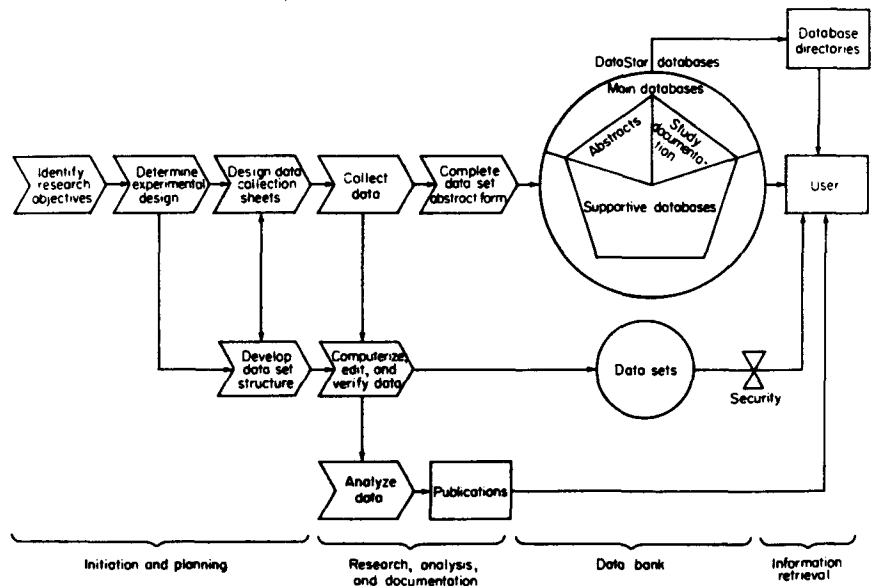
Price: $21

Order Your Copy Now

Figure 1. Integration of the Forest Science Data Bank in the research process—from the initial stages of research planning, through analysis and data documentation, to subsequent information retrieval.

Inc., 1980), a general-purpose, relational database management system. The FSDB is unique in that information on study documentation is entered into several integrated DataStar databases: two main databases—Abstracts and Detailed Study Documentation—and six supportive databases—References, Contact Persons, Computer Programs, Variable Formats, Variable Codes, and Data File Organization *(fig. 1)*.

Each DataStar database comprises a series of records. For every data set, one record is entered into Abstracts and a complementary record into Detailed Study Documentation. Abstracts primarily contain information describing the data set for search and retrieval; thus, the uninformed user can search all documented data sets with key words such as "investigators," "research locations," or "tree species."

Detailed Study Documentation contains information summarizing a study's purpose, design, and methodology, and describes supporting documents and their locations. The supportive DataStar databases differ from the two main databases in that each record contains information (such as citations and addresses of investigators) potentially applicable to several data sets.

To capture the needed information, we developed a series of standard forms (copies available on request from the senior author). Much of the information required by these forms is generated by a researcher regardless of the data bank—a selling point of the system to our users. A standard set of forms ensures uniform documentation for each data set, ready for entry into the data bank.

Our most challenging problem now is to gain the full cooperation of re-

searchers in providing complete documentation of existing data sets. Only their cooperation will ensure that our information is current and that standards of documentation and software development best serve the needs of all users. ∎

Literature Cited

MICROPRO INTERNATIONAL, INC. 1980. DataStar™ software. Micropro International Corp., San Rafael, CA.

WAGNER, R.G. 1981. Computerized data form user's guide for the CRAFTS B-level release studies. Unpubl. manuscr. Dep. For. Sci., Oreg. State Univ., Corvallis.

Susan G. Stafford is assistant professor, Paul B. Alaback is research associate, Greg J. Koerper is a graduate student, and Mark W. Klopsch is research assistant, Department of Forest Science, Oregon State University, Corvallis 97331. This is paper FRL 1618, Forest Research Laboratory, Oregon State University.