

1979

Biobank, a Computerized Data Storage and Processing System for the Vascular Flora of Iowa

Lawrence J. Eilers
University of Northern Iowa

Copyright © Copyright 1979 by the Iowa Academy of Science, Inc.
Follow this and additional works at: <http://scholarworks.uni.edu/pias>

Recommended Citation

Eilers, Lawrence J. (1979) "Biobank, a Computerized Data Storage and Processing System for the Vascular Flora of Iowa," *Proceedings of the Iowa Academy of Science*: Vol. 86: No. 1 , Article 6.
Available at: <http://scholarworks.uni.edu/pias/vol86/iss1/6>

This Research is brought to you for free and open access by UNI ScholarWorks. It has been accepted for inclusion in Proceedings of the Iowa Academy of Science by an authorized editor of UNI ScholarWorks. For more information, please contact scholarworks@uni.edu.

Biobank, a Computerized Data Storage and Processing System for the Vascular Flora of Iowa

LAWRENCE J. EILERS

Department of Biology, University of Northern Iowa, Cedar Falls, Iowa 50613

BIOBANK is an electronic data processing (EDP) system for storing and processing the accumulated data on the vascular flora of Iowa. This computerized system was designed as a tool to facilitate floristic research by systematists who may have had little or no training in computer operation. Thus, a minimum of coding is used; the usual kinds of information are entered into the system and the printouts are in traditional formats. The programs, which together form the BIOBANK system, are written in the PL/1 programming language. The data processed by the system can be: 1) of varying lengths; 2) stored in the data bank; 3) modified or corrected; 4) merged as desired; and 5) retrieved in a variety of ways by the search and store facilities of the system. Because of the flexibility of the BIOBANK system, it can be readily adapted to process other kinds of data.

INDEX DESCRIPTORS: EDP system for floristic data, computerization of Iowa flora, Iowa vascular flora, BIOBANK EDP system.

A vast amount of data on the Iowa vascular flora exists in the literature (summaries in Peck, 1976a and 1976b, and in Eilers, 1975), in the several Iowa herbaria, and elsewhere. It was clear early in the flora of Iowa project that a large part of the work would consist of bringing these data together in an ordered, usable form. As a consequence, it was necessary to decide at that time what kinds of data to gather and in what form. A further consideration was the possible end uses that could be made of these data. If a simple ordered listing of the data would be sufficient for future needs, then the time-honored system of hand-written or typed data cards would be adequate for the job. However, if one is working with large quantities of data and wishes to sort the data into different arrangements, to search for particular kinds of data items (singly, or in combinations) or to print out the data quickly in a variety of formats, then a computerized electronic data processing system (EDP) has distinct advantages. Deciding in favor of an EDP system with its desirable features, I began to search for an existing system that could be adapted to my needs.

The system that seemed most promising was written in the PL/1 language and developed by Dr. Theodore Crovello and his associates at the University of Notre Dame, Notre Dame, Indiana (Crovello, 1972). Dr. Crovello graciously made the programs available for use, but after a trial period the conclusion was reached that it would require an extensive modification of the Notre Dame system to fit it to my particular needs. Thus, I designed my own computerized system, wrote the necessary programs in PL/1 language and tested them extensively. The result is the BIOBANK EDP system designed particularly for the storage and processing of the systematic data on the Iowa vascular flora.

DESIGN CRITERIA

In all aspects of the design of the BIOBANK system, much effort was put into what is frequently termed "human engineering": designing for easy use of the system by biologists who may have little knowledge of computer programming and operation. Thus, the input data are nearly all in traditional word, sentence, or number form, and can be easily read and understood. The edited printouts are in familiar textual formats.

Other criteria used in developing the system are as follows:

- 1) card input for broader usability,
- 2) hierarchical arrangement of taxonomic data,
- 3) uncoded data (to the extent possible),
- 4) storage and processing of both alphabetic and numeric data,
- 5) handling of data of varying lengths,
- 6) facilities for searching and/or sorting the data as desired,

- 7) an expandable system,
- 8) a flexible system,
- 9) ease of updating the sorted data: correcting errors, adding data, and deleting data,
- 10) eventual use for computerized mapping.

Because of the effort spent on the design, and because of the versatility of the PL/1 programming language, all of these features are now incorporated into the BIOBANK system.

THE DATA RECORDS

All data are entered and stored in BIOBANK in record format (Figs. 1 and 2). Each record is made up of a taxon code and the taxon name, followed by the data items (fields) which describe the given taxon or specimen. These fields are prefixed with a field number and delimited by slashes and are arranged end-to-end in a prescribed order, usually the order used in the literature or that used for specimen labels. One record contains the data items for one taxon (family, genus, or species) (Fig. 1), or one specimen (Fig. 2). The records are of varying lengths, depending on the amount of information available.

THE SYSTEM

The BIOBANK system (Fig. 3) is composed of the stored data in record format and seven integrated computer programs that serve the following data-processing functions:

- 1) prechecking records for format errors before storage,
- 2) storing records tightly packed in the data bank,
- 3) updating stored records,
- 4) searching for particular data items (binomials, habitats, localities, etc.) either singly or in combinations, with a printout of the records containing the desired items,
- 5) sorting and printout of records according to an alphabetical or numeric ordering of the data items within the records,
- 6) merging taxon data records and their associated specimen data records with a subsequent printout of the results,
- 7) printing of specimen labels, catalogs, text, etc., in punctuated, edited form (Figs. 4 and 5).

For convenience, the data records are stored in BIOBANK alphabetically by the formal taxon names, though they can be retrieved in various other orders by the sort and search facilities of the system. A printout of the records in systematic order (e.g., the Dalla Torre and Harms sequence) could be accomplished by the simple addition of a

```

a01.0CUP --- --- 00/02.0CUPRESSACEAE/02.2CYPRESS F
AMILYS$

a01.0CUP JUN --- 00/03.0JUNIPERUS L./03.2JUNIPER G
ENUS$

a01.0CUP JUN COM 00/04.0JUNIPERUS COMMUNIS L./04.3
COMMON JUNIPER/05.0DRY WOODED BLUFFS, ROCKY SLOPES
/06.0RARE NORTHEAST, FREQUENT EXTREME NORTHEAST/07
.0RARE$

a01.0CUP JUN VIR 00/04.0JUNIPERUS VIRGINIANA L./04
.3EASTERN RED CEDAR/05.0DRY, OPEN WOODS, ROCKY BLU
FFS, PASTURES/06.0COMMON EASTERN HALF, FREQUENT SO
UTHWEST, INFREQUENT NORTHWEST/07.0FREQUENTS$

a01.0PIN --- --- 00/02.0PINACEAE/02.2PINE FAMILY$

a01.0PIN ABI --- 00/03.0ABIES MILL./03.2FIR GENUS$

a01.0PIN ABI BAL 00/04.0ABIES BALSAMEA (L.) MILL./
04.3BALSAM FIR/05.0STEEP, NORTH-FACING BLUFFS/06.0
RARE IN EXTREME NORTHEAST/07.0RARE$

a01.0PIN PIN --- 00/03.0PINUS L./03.2PINE GENUS$

a01.0PIN PIN STR 00/04.0PINUS STROBUS L./04.3WHITE
PINE/05.0STEEP, USUALLY SANDY WOODED SLOPES AND L
EDGES/06.0RARE EASTERN HALF, FREQUENT EXTREME NORT
HEAST/07.0INFREQUENTS$

a01.0TAX --- --- 00/02.0TAXACEAE/02.2YEW FAMILY$

a01.0TAX TAX --- 00/03.0TAXUS L./03.2YEW GENUS $

a01.0TAX TAX CAN 00/04.0TAXUS CANADENSIS MARSH./04
.3AMERICAN YEW/05.0STEEP, MOIST, CALCAREOUS WOODED
SLOPES, TALUS/06.0INFREQUENT TO FREQUENT NORTHEAS
T, RARE SOUTHEAST/07.0INFREQUENTS$
    
```

Fig. 1. *Gymnosperm taxon records as stored in BIOBANK. The 1st record is the family record for the Cupressaceae, the 2nd is the genus record for Juniperus, and the 3rd and 4th records contain the data for the 2 Iowa species of Juniperus. See Fig. 4 for an edited printout of these records.*

taxon number to each record and then sorting the records by these numbers.

The BIOBANK system is implemented in the batch mode on the IBM 360/65 computer at the University Computing center (UCC), University of Iowa, Iowa City. Thus far we have been limited to card input at our connecting terminal at the University of Northern Iowa, but we will soon have in operation an interactive terminal that we can use to make tapes of the input data. We will then feed these tapes into the computer at Iowa City. This method of data entry will allow us to pre-edit the data more carefully and to bypass the hand-punching of cards with its higher error rate. Cards for a back-up data file can then be punched by the computer.

```

a01.0AST KRI DAN 01/04.0KRIGIA DANDELION (L.) NUTT
./26.0USA: IOWA:/27.0UNION COUNTY:/28.2CRESTON/31.
0T.L. ANDREWS/32.0ISC22350/35.0KRIGIA VIRGINICA/35
.1K. DANDELION (L.) NUTT., R. A. DAVIUSON,1952 04
11$

a01.0AST SEN VUL 01/04.0SENECIO VULGARIS L./26.0US
A: IOWA:/27.0BOONE COUNTY:/28.2HWY. 30, E. EDGE OF
BOONE./29.0FRONT DOOR OF BOONE RAMBLER AGENCY/30.
0ONE PLANT, NEW TO IOWA./31.0R. W. POHL/31.19361/3
2.0ISC236248/33.01963 06 14$

a01.0BET BET PUM 01/04.0BETULA PUMILA L. VAR. GLAN
DULIFERA REGEL/26.0USA: IOWA:/27.0ALLAMAKEE COUNTY
:/28.2POSTVILLE/31.0ORVILLE SCHULTZ/32.0ISC84200/3
3.01913 08 00$

a01.0BET BET PUM 02/04.0BETULA PUMILA L. VAR. GLAN
DULIFERA REGEL/26.0USA: IOWA:/27.0CHICKASAW COUNTY
:/28.2OUTSKIRT OF NEW HAMPTON/29.0HOG/31.0MARGARET
MURLEY/31.11549/32.0ISC160390/33.01940 09 13$

a01.0FAB BAP AUS 01/04.0BAPTISIA AUSTRALIS (L.) R.
BR. VAR. MINOR (LEHM.) WATS./26.0USA: IOWA:/27.0H
OWARD COUNTY:/28.0SEC. 33, T100N-R13W/28.2HAYDEN P
RAIRIE/31.0R. D. ALBERTSON/31.1304-99/32.0ISC27908
1/33.01969 08 06$

a01.0FAB BAP AUS 02/04.0BAPTISIA AUSTRALIS (L.) R.
BR. VAR. MINOR (LEHM.) WATS./26.0USA: IOWA:/27.0D
ECATUR COUNTY:/28.2IOWA #2, 1.7 MILES EAST INTERSE
CTION WITH STATE #294/31.0DUANE ISELY AND DAN NIFF
ENEGGER/31.19812/32.0ISC248616/33.01965 08 31/35.0
BAPTISIA MINOR LEHM.$
    
```

Fig. 2. *Angiosperm specimen records as stored in BIOBANK. See Fig. 5 for an edited printout of these records.*

PRESENT STATUS OF THE PROJECT

The BIOBANK system has been thoroughly tested and is in operation. All taxon data records for the pteridophytes and gymnosperms of Iowa are now in computer storage, as are the taxon and specimen records for a number of threatened or endangered angiosperm species. An initial working list of angiosperm species records has been punched and checked by the system for format errors. The next step is to edit the records visually before entering them in the system. We have also begun to extract data from the specimens in the Martin Grant Herbarium at UNI for computer entry.

LIMITATIONS AND ADVANTAGES OF THE BIOBANK SYSTEM

It should be made clear that the BIOBANK system does not eliminate the need for careful, scholarly floristic work. One must still search the literature, study specimens, and evaluate the validity of the data. Also, nomenclatural decisions still have to be made; the problems of synonymy still have to be worked out, and the use of terminology has to be standardized. Thus, the BIOBANK system does not reduce the time and effort spent in studying and evaluating the data. The advantages lie in the greatly increased speed of data handling and the expanded facilities for processing the data once it has been stored in the computer.

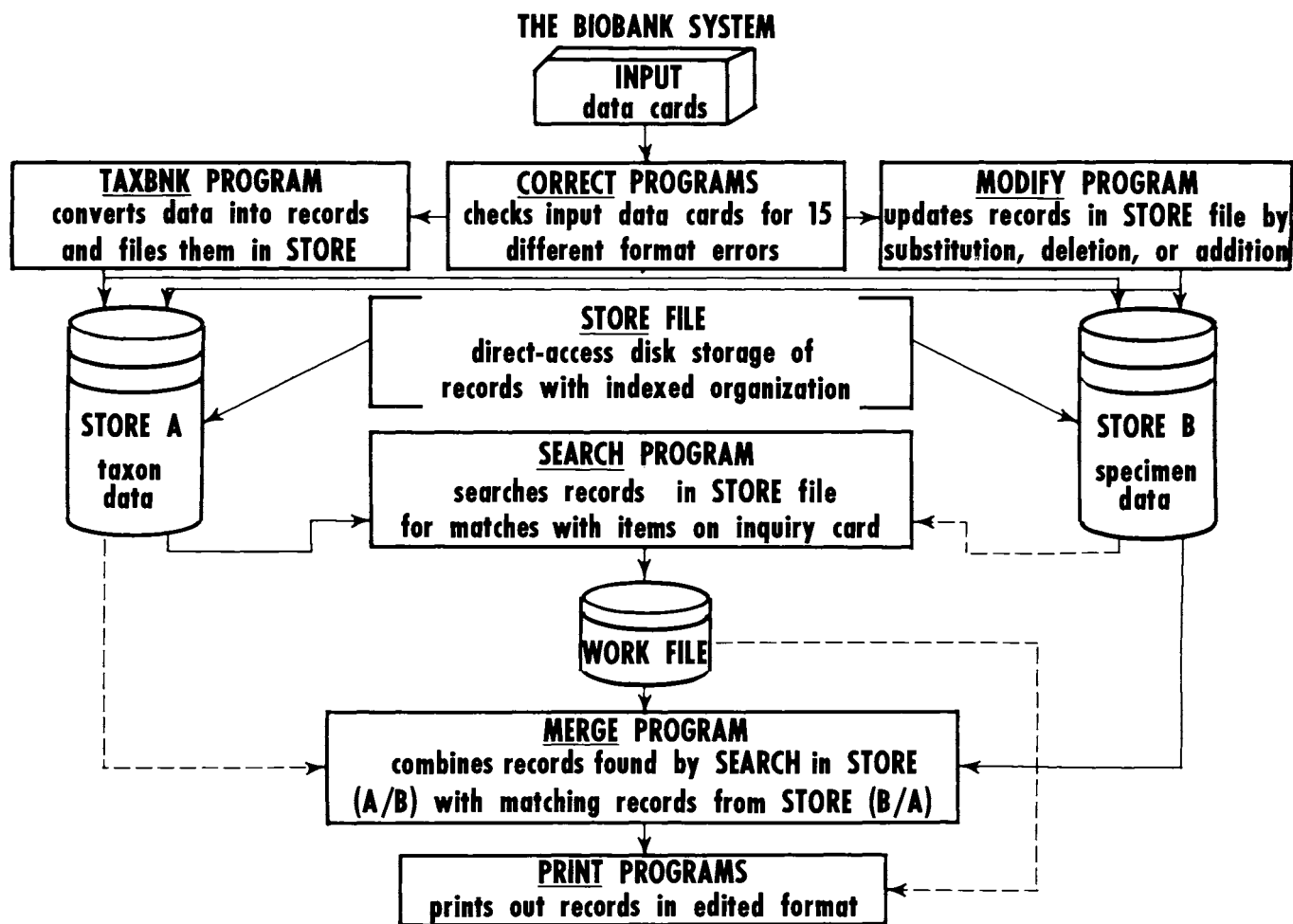


Fig. 3. The BIOBANK system showing its storage facilities, the functions of the computer programs that process the data records, and the interrelationships of these components.

The SEARCH program, for example, can look for the particular records that do or do not contain one or all of several data items, and the selected records are printed out within a brief period of time. For example, one could ask the system to print out a list of all the known *flowering plants* that bloom in *April, May, or June* in the *upland woods* (but not the *alluvial woods*) of *Black Hawk County*. One could also ask for a flora of any given county, a list of all counties in which a particular species is known to occur, the species associated with a given type of habitat, or a list of all plants not reported in Iowa for the past 50 years. Furthermore, the listing generated by any search procedure can be sorted into either alphabetic or numeric order before printing it out. The above examples illustrate just a few of the many ways in which the facilities of the BIOBANK system can be utilized to generate useful information.

CONCLUSIONS

The BIOBANK system was designed to be a tool to facilitate research on the Iowa vascular flora. Much effort was expended in making it easy to use with traditional kinds of systematic data but with the

advantages derived from computerized data processing. The data entered are recognizable and the output is familiar to systematists. Also, the information stored in BIOBANK can be changed as needed.

I invite the assistance of other workers on the Iowa vascular flora. A computerized state flora will serve many needs, and the more workers we have, the faster the flora will be completed. I also hope that the BIOBANK system will have utility for researchers working with other groups of organisms. Because of the flexibility designed into the BIOBANK system it can be easily adapted to process many other kinds of information. For additional information on the BIOBANK system and its applications, please feel free to contact me.

ACKNOWLEDGEMENTS

Work on this project was supported in part by a Faculty Research Fellowship awarded for summer of 1977 by the Graduate College of the University of Northern Iowa. I am very grateful for this assistance. I am also indebted to the College of Natural Sciences at UNI for funds for card punching and computer operation, the members of the UNI Academic Computing Services for many kinds of assistance given freely

CUPRESSACEAE, CYPRESS FAMILY

JUNIPERUS L., JUNIPER GENUS

JUNIPERUS COMMUNIS L., COMMON JUNIPER.
 DRY WOODED BLUFFS, ROCKY SLOPES; RARE NORTHEAST,
 FREQUENT EXTREME NORTHEAST; RARE.

JUNIPERUS VIRGINIANA L., EASTERN RED CEDAR.
 DRY, OPEN WOODS, ROCKY BLUFFS, PASTURES; COMMON
 EASTERN HALF, FREQUENT SOUTHWEST, INFREQUENT
 NORTHWEST; FREQUENT.

PINACEAE, PINE FAMILY

ABIES MILL., FIR GENUS

ABIES BALSAMEA (L.) MILL., BALSAM FIR.
 STEEP, NORTH-FACING BLUFFS; RARE IN EXTREME
 NORTHEAST; RARE.

PINUS L., PINE GENUS

PINUS STROBUS L., WHITE PINE.
 STEEP, USUALLY SANDY WOODED SLOPES AND LEDGES;
 RARE EASTERN HALF, FREQUENT EXTREME NORTHEAST;
 INFREQUENT.

TAXACEAE, YEW FAMILY

TAXUS L., YEW GENUS

TAXUS CANADENSIS MARSH., AMERICAN YEW.
 STEEP, MOIST, CALCAREOUS WOODED SLOPES, TALUS;
 INFREQUENT TO FREQUENT NORTHEAST, RARE SOUTHEAST;
 INFREQUENT.

KRIGIA DANDELION (L.) NUTT.

USA; IOWA: UNION COUNTY: CRESTON; T.L. ANDREWS;
 ISC22350.
 FIRST IDENT. AS: KRIGIA VIRGINICA
 FIRST ANNOTATION: K. DANDELION (L.) NUTT., R. A.
 DAVIDSON, 1952 04 11

SENECIO VULGARIS L.

USA; IOWA: BOONE COUNTY: HWY. 30, E. EDGE OF
 BOONE.; FRONT DJOR OF BOONE RAMBLER AGENCY; ONE
 PLANT, NEW TO IOWA.; R. W. POHL, 9361; ISC236248;
 1963 06 14.

BETULA PUMILA L. VAR. GLANDULIFERA REGEL

USA; IOWA: ALLAMAKEE COUNTY: POSTVILLE; ORVILLE
 SCHULTZ; ISC84200; 1913 08 00.

USA; IOWA: CHICKASAW COUNTY: OUTSKIRT OF NEW
 HAMPTON; BOG; MARGARET MURLEY, 1549; ISC160390;
 1940 09 13.

BAPTISIA AUSTRALIS (L.) R. BR. VAR. MINOR (LEHM.)
 WATS.

USA; IOWA: HOWARD COUNTY: SEC. 33, T100N-R13W;
 HAYDEN PRAIRIE; R. D. ALBERTSON, 304-99;
 ISC279081; 1969 08 06.

USA; IOWA: DECATUR COUNTY: IOWA #2, 1.7 MILES EAST
 INTERSECTION WITH STATE #294; DUANE ISELY AND DAN
 NIFFENEGGER, 9812; ISC248616; 1965 08 31.
 FIRST IDENT. AS: BAPTISIA MINOR LEHM.

Fig. 4. Edited and punctuated printout of the taxon data records in Fig. 1.

Fig. 5. Edited and punctuated printout of the specimen data records in Fig. 2.

and cheerfully and to the Department of Biology for released time for this project and for research assistants. I wish to thank Dr. Theodore J. Crovello, Department of Biology at Notre Dame University, for the assistance noted earlier in the text of this paper. I am especially grateful to Dr. John C. Downey, Head of the Department of Biology, who suggested computerizing the data on the Iowa vascular flora, and who gave me unflinching encouragement and support.

REFERENCES

CROVELLO, T.J. 1972. Computerization of specimen data from the Edward Lee Greene Herbarium (ND-G), at Notre Dame. *Brittonia* 24:131-141.

EILERS, L.J. 1975. History of studies on the Iowa vascular flora. *Proc. Iowa Acad. Sci.* 82:59-64.

PECK, J.H. 1976a. An annotated bibliography to the literature on pteridophytes in Iowa. *Proc. Iowa Acad. Sci.* 82:203-208.

PECK, J.H. 1976b. The pteridophyte flora of Iowa. *Proc. Iowa Acad. Sci.* 83:143-160.