Hunt Institute for Botanical Documentation
5th Floor, Hunt Library
Carnegie Mellon University
4909 Frew Street
Pittsburgh, PA 15213-3890
Telephone: 412-268-2434
Email: huntinst@andrew.cmu.edu
Web site: www.huntbotanical.org

The Hunt Institute is committed to making its collections accessible for research. We are pleased to offer this digitized item.

Usage guidelines
We have provided this low-resolution, digitized version for research purposes. To inquire about publishing any images from this item, please contact the Institute.

Statement on harmful and offensive content
The Hunt Institute Archives contains hundreds of thousands of pages of historical content, writing and images, created by thousands of individuals connected to the botanical sciences. Due to the wide range of time and social context in which these materials were created, some of the collections contain material that reflect outdated, biased, offensive and possibly violent views, opinions and actions. The Hunt Institute for Botanical Documentation does not endorse the views expressed in these materials, which are inconsistent with our dedication to creating an inclusive, accessible and anti-discriminatory research environment. Archival records are historical documents, and the Hunt Institute keeps such records unaltered to maintain their integrity and to foster accountability for the actions and views of the collections' creators.

Many of the historical collections in the Hunt Institute Archives contain personal correspondence, notes, recollections and opinions, which may contain language, ideas or stereotypes that are offensive or harmful to others. These collections are maintained as records of the individuals involved and do not reflect the views or values of the Hunt Institute for Botanical Documentation or those of Carnegie Mellon University.

About the Institute
The Hunt Institute for Botanical Documentation, a research division of Carnegie Mellon University, specializes in the history of botany and all aspects of plant science and serves the international scientific community through research and documentation. To this end, the Institute acquires and maintains authoritative collections of books, plant images, manuscripts, portraits and data files, and provides publications and other modes of information service. The Institute meets the reference needs of botanists, biologists, historians, conservationists, librarians, bibliographers and the public at large, especially those concerned with any aspect of the North American flora.

Hunt Institute was dedicated in 1961 as the Rachel McMasters Miller Hunt Botanical Library, an international center for bibliographical research and service in the interests of botany and horticulture, as well as a center for the study of all aspects of the history of the plant sciences. By 1971 the Library's activities had so diversified that the name was changed to Hunt Institute for Botanical Documentation. Growth in collections and research projects led to the establishment of four programmatic departments: Archives, Art, Bibliography and the Library.

May 1, 1969

Summary of The Botanical Type Collection and Dissemination System (BOTCAD)

Purpose of the System

An automated information collection and dissemination system is being developed to create a data bank containing "type" data derived from botanical specimens throughout the world. A record for each "type specimen" is prepared that contains:

1) historical facts concerning the specimen and

2) a list showing the institution acronym, catalog number, and kind of type reported by each herbarium for duplicates of the specimen.

Each record, printed on a 5 by 8 index card, is also made available on magnetic tape. A replication of the data bank will be furnished to each institution that participates in the project.

Methods

Data, derived from botanical types at the NMNH are entered into the system on a continuing basis. Records representing selected genera are printed on index cards which are mailed to participating institutions. A two card data record, perforated for easy burst, is furnished for each species. The first card is a "Record of Holdings" which the recipient may wish to keep on file until it is superseded by a more comprehensive record. On the first distribution, the cards may contain only data from the NMNH. Subsequent distributions will reflect the holdings of all participating institutions that have reported a kind of type for that taxon.

The second card is a "Data Collection Form". If the recipient has in his herbarium a duplicate of the specimen cited on the first card, he is asked to complete line 12 of the second card (Institution acronym,

Digitized by the Hunt Institute for Botanical Documentation

catalog number, and kind of type), and to detach and return the card. If the recipient has other verified types, especially of taxa recently described, he may enter the data by completing lines 1 through 12.

"Data Collection Forms" returned to the Smithsonian will cause the system to update any type already on record and to incorporate any reported type that is new to the file. All data thus reported is reflected in subsequent distributions for the taxon.

Management Information Reports

The system automatically adjusts its activities according to the status of records in its data bank and provides the following management information:

1) Genera activated for a dissemination-collection cycle.

2) Genera fully updated and disseminated and consequently deactivated.

3) Institutions that have been furnished current systems data for selected genera and asked to respond.

4) System's holdings by family, genus, etc.

5) Each institution's holdings by family, genus, etc.

6) Other miscellaneous items of information.

The system prevents redundant mailing of records (unless otherwise directed) and insures that no institution in the network is by-passed regarding any group, when the various herbaria report randomly on genera of their choice at differing points in time. The system rapidly processes large volumes of data; facilitates data validation and correction; and prints accurate and timely reports on the collective holdings of herbaria throughout the world.

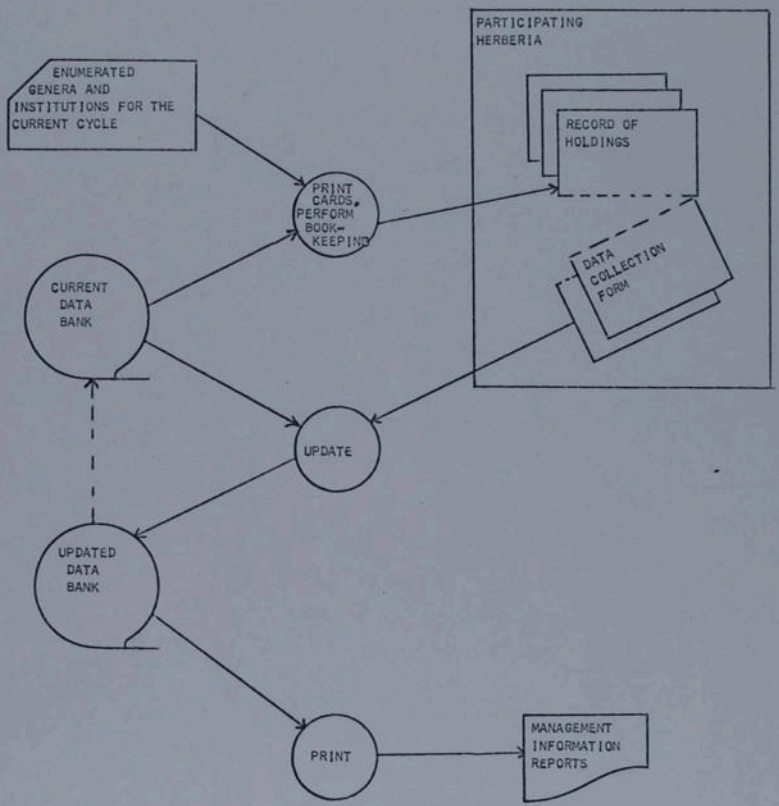
The stringent housekeeping and bookkeeping chores involved in a comprehensive information collection and dissemination system are in large measure handled by the computer.

Digitized by the Hunt Institute for Botanical Documentation

April 30, 1969, ISD

Botanical Types

Schematic of Data Collection and Dissemination



Digitized by the Hunt Institute for Botanical Documentation

TYPE REGISTER - CATALOG CARD

SCROPHULARIACEAE

MIMULUS DISCOLOR GRANT

ANN. MO. BOT. GARD. 11:257. 1924.

COLLECTORS: PURPUS, C. A. COLL. NO.: 5311

DATE COLL.: APR 1897

US; CALIFORNIA; PAH UTE PEAK; -

US...0328682 ISOTYPE

THIS CARD MAY BE RETAINED FOR YOUR FILES

RECORD ID. 314

TYPE REGISTER - DATA COLLECTION CARD. 73905645178381219731

IF YOU HAVE IN YOUR HERBARIUM A DUPLICATE OF THE ABOVE SPECIMEN, COMPLETE LINE 12 BELOW ON THIS CARD.

IF YOU HAVE OTHER VERIFIED TYPES, ESPECIALLY OF TAXA WHICH YOU HAVE RECENTLY DESCRIBED, PLEASE ENTER DATA AS COMPLETELY AS POSSIBLE ON THIS OR OTHER CARDS AS NEEDED. RETURN TO DEPARTMENT OF BOTANY, SMITHSONIAN INSTITUTION, WASHINGTON, D.C. 20560.

01. FAMILY: RECORD ID. 000314
02. GENUS: SPECIES:

03. VAR. OR SUBSP.: FORMA:

04. AUTHOR:
05. CITATION:

06. COLLECTORS: COLL. NO.:
COLL. DATE:

07. LOCALITY: COUNTRY:
08. STATE OR PROVINCE:
09. COUNTY OR DEPARTMENT:
10. TOWN OR SUBREFERENCE:

11. REMARKS:

12. INSTITUTION ACRONYM: CAT. NO: KIND OF TYPE:

Digitized by the Hunt Institute for Botanical Documentation

DATA ON LABELS

GENERAL CONSIDERATIONS

Before a program for the storage and retrieval of data is written, it is important to make it as compatible as possible with the idiosyncrasies in the documentation of past and present herbarium label formats. The chief difficulty in developing an efficient system for electronic data storage is the incompleteness of available records. Many herbarium specimen labels suffer in this respect in one or more items of information. In addition, the various parts of an individual record should not be too long and should not have to be coded before input either by the botanist or by the machine operator. Coding adds further chance for errors and takes time from the input operation. Some types of coding can be handled by the computer, such as the coding of scientific names by means of a "taxonomic dictionary."

Because the working up of the "backlog" of data on available specimen labels will be hindered by imcompleteness and illegibility of the labels, it is important that efforts be made as soon as possible to improve the information which is becoming available on all the new specimens being added to herbaria day by day. Therefore, we have considered in detail the data associated with herbarium labels from the various angles which have a bearing on the development of a computerized information retrieval system. The items discussed below are mostly marked to indicate into which of three categories we believe they fall; some we consider essential (E), others desirable (D) and the rest optional (O).

LABEL FORMAT

The dimensions of the label do not need to be standardized but in the case of bryophytes and lichens stored in packets or envelopes, the label size should not exceed that of the container. Normally there will be no limit to the choice of format for the label. The order in which the information is arranged may be restricted if the labels are produced as a byproduct of recording data on a preprinted catalogue record form. Each herbarium would be free to determine its own label format and recording system while still able to participate in a nation-wide data storage project.

PLANT NAME

(E) The label must bear the scientific name of the taxon with at least the binomial consisting of the generic epithet, specific epithet and author name(s). The name should appear near the top of the label where it may be seen easily.

(D) If the identification has been carried to a lower level than species, it should be so stated by adding the epithets and author name(s) down to the lowest level, using ssp, for subspecies, var. for variety and forma for form. Author name(s) are omitted for infraspecific taxa which represent the typical element of the taxon and repeat the preceding epithet.

(O) Since specimens with imcomplete identification may constitute valuable records in a data system, it is acceptable to give the name of the genus followed by "sp." which signifies "species not determined".

O

Digitized by the Hunt Institute for Botanical Documentation

Orthography. - No blank spaces should occur within an epithet. Special abbreviation such as "cf.", "aff.", "prope" or "?" should not be used between the generic and specific epithets but comments of this type may be placed under "Notes". Abbreviations for authors' names should be consistent with the usage in an up-to-date flora or manual.

Synonymy. - There will always be disagreement about the hierarchical status of some plant names and also there will be in use a certain amount of valid synonymy. It is hoped that the C.B.A. checklists now in preparation will serve as guides to synonymy. If synonyms are to be shown at all on labels, they should be given under "Notes".

Hybrids. - Interspecific hybrids should be described by the single name (if available) without multiplication symbol (X) after the generic epithet or, if necessary, by the scientific names of the two parents separated by X.

Types. - If a specimen constitutes TYPE material, the scientific name should be followed by the statement "TYPE" and more particular information about the kind of TYPE (ISOTYPE, HOLOTYPE, etc.) may be put under "Notes". Definitions for kinds of TYPES may be found in McVaugh et al., 1968.

Numerical Coding. - The coding of a taxon by number is not necessary but, if desired, it may be used in local information systems to develop a more rapid input of specimen data or sight records. It is also useful in a simple system based on tab cards where economy of space precludes the use of the full scientific name.

LOCATION

(E) A locality name (word reference) and a grid location (numerical coordinates) must be given on the label. The locality name should be an accepted geographical name of a settlement or a physiographic feature as listed in one of the official gazetteers. Distance (in kilometres or miles) and direction (compass bearing or cardinal point) from the place named are optional statements. If a local name is not specific (e.g. "Gull Lake") then the township and county should be added. The numerical reference will help pinpoint the location when the local name does not do so.

(E) The grid location must be one of the three indications, namely the latitude and longitude to the nearest minute, the U.T.M. (Universal Transverse Mercator) grid reference to the nearest 10 X 10 km square, or the 1:5,000 map sheet number followed by the U.T.M. reference to the 10 X 10 km unit by means of two digits.

(O) The label may also contain some descriptive regional or local names. The first is often given in the title, e.g. Flora of the Maritimes, Plants of the Bruce Peninsula, etc. The name of the Province or Territory, County or District, National or Provincial Park or title of a physiographic or vegetational region may be of interest in floristic studies.

Altitude. - The altitude may not be of great importance in lowlands, but for mountainous regions it becomes an essential part of the label. Unless a special study is being made, an estimate to the nearest hundred metres is usually sufficient.

Digitized by the Hunt Institute for Botanical Documentation

Cultivated material. - When a specimen has been collected in a field plot, test garden, greenhouse or growth chamber, no location should be given in the space ususally allocated for this information. The statement "Cultivated - see Notes" should be inserted and the details given under "Notes."

HABITAT

(D) Some statement of the nature of the environment in which the specimen was collected is important in many cases. A standardization of this type of information does not seem feasible. A short characterization might stress either the topographical or pedological aspects or a specific factor complex such as the type of human disturbance or the dominant life form of the habitat. Lists of dominant plants or associated species should be relegated to the notes but the name of the host plant for an epiphyte or a parasite should be given as part of the habitat description.

DATE OF COLLECTION

(E) The day, month and year of the collection must be stated and in that particular order.

(O) To simplify input procedures, dates on catalogue cards or on tab cards and paper tapes may be given in arabic numerals separated by slashes, e.g. 5/7/1968. There are sufficient old collections to necessitate the use of more than the last two digits for the year.

COLLECTOR & COLLECTOR'S NUMBER

(E) The label must contain the initials and surname of the principal collector of the specimen. If more than one person is listed, the person whose number is given must be cited first.

(O) Titles such as Mrs., Prof., Col., Dr. etc. should be omitted from labels. If the collector is, for example, a Mrs. M.I. Smith, then the initials must be her own and not those of her husband, otherwise there is no distinction when the title is dropped.

Names of additional collectors should be limited to two and the statement J.H. Smith et al. is of little value on the label itself. It has its proper place in abbreviating a list of three or more names when information from a label is cited in a publication.

(D) Each collection (except for exact duplicates in time and space) should receive a unique number allotted by each collector in chronological sequence. Collectors should not begin again at number 1 each new season. Numbers involving coding for geographical regions (e.g. 529AG for a specimen collected at Arkona Glen) or including dates (e.g. 68529 for number 529 collected in 1968) should be avoided. Collector's numbers should not contain blanks, dashes, slashes or any other symbols. It is acceptable, when necessary to designate divided material resulting from accidental mixed collections, to append a single alphabetic character to the number (e.g. 12529A and 12529B) for two different taxa originally assigned the number 12529).

HERBARIUM CODE & ACCESSION NUMBER

(E) Each herbarium sheet is normally stamped with the distinctive seal or stamp of the herbarium to which it belongs. The internationally accepted codes (e.g. CAN, OK, SASK. TRT, etc.) are ideal for use in data processing and these codes are usually not more than four letters in length.

Digitized by the Hunt Institute for Botanical Documentation

(O) In order to designate a particular sheet in a herbarium, it is necessary that each sheet be stamped with a unique number in a single series for that herbarium. This number is the sheet number or an accession number and bears no relation to collector's numbers. It is not sufficient to locate a specimen by the collector's number because so-called "duplicates" have been known to e a different taxon or to contain a mixture. More important, however, is the fact that many collectors have never used collector's numbers and others have repeated their numbering sequence in successive seasons.

DETERMINATOR

(D) The label should bear the name (initials and surname) of the person who identified the specimen, preceded by "Det." and followed by the year in four digits. Any subsequent change of name should be made on a standard annotation label glued to the sheet near the original label. All such annotations should be made in ink or by means of printed correction slips and should bear the name of the revisor and the date.

NOTES

(D) It is useful to note near the bottom of the label the number of replicas prepared from the same collection. The label may also allow a space of two or three lines at the bottom, possibly with a printed heading "Notes", where additional information can be given. Statements about the height of the plant, colour of flowers, abundance, local names or uses, references to photographs, chromosome counts or samples of seeds taken can be recorded here. In a data bank it may not be feasible to search this information for specific features, but the Notes can be stored and recalled in toto.

The name and number of an exsiccata specimen should be entered under notes and the title of an expedition may also be useful information.

In the case of TYPE material, bibliographic reference for the new taxon should give the place and date of publication for ready reference.

-- Extract from report by R.E.Beschel & J.H.Soper, [illegible], [illegible]
March 31, 1969, Committee on Automation and
Standardization of Herbarium Procedures,
Canadian Botanical Association.

Digitized by the Hunt Institute for Botanical Documentation

Table 2. Problems in the use of EDP in biological collections.

A. Recognition of those uses of EDP that have scientific value.

B. Physical
   1. Existence of necessary equipment (hardware and software)
   2. Availability of equipment to curators
   3. Compatibility of an EDP system with different systems at other collections
   4. Availability of space (for equipment and for data)

C. Cost
   1. Price of equipment
   2. Price of trained EDP personnel
   3. Source of funds for EDP
   4. Price of time demanded of the curator to initiate and maintain an error-free system of information

D. Human
   1. Fear of EDP by curators
   2. Ignorance of EDP by curators
   3. Inertia generated by awe at magnitude of certain EDP tasks

E. The collection
   1. Adequacy of the collection to produce reliable results
   2. Specimen misidentification
   3. Synonymy
   4. Choice of the particular specimens to process

F. Organizational level of institutions for studies using EDP

G. Optimal level of taxonomists organization for studies using EDP

H. Generation of a suitable context to decide if use of EDP is worth the expenditure of the required resources in a particular situation

I. Codification of biological names

Digitized by the Hunt Institute for Botanical Documentation

INFORMAL CONFERENCE ON DATA PROCESSING
IN THE HERBARIUM

DEPARTMENT OF BOTANY, SMITHSONIAN INSTITUTION
May 5, 1969 - 9:30 a.m.

Suggested Morning Program: (one-half hour per speaker with discussion)

Soper: Data processing for current herbarium acquisitions

Rogers: Data processing of herbarium holdings

Crovello: Data processing of a specialized herbarium collection

Hale: Analysis of the Department of Botany Type Register

Shetler: Systems compatibility

Suggested Afternoon Program:

Discussion of the following questions:

What kinds of questions can we expect from EDP users?

How can a large herbarium with limited resources utilize EDP methods?

How much coordination of efforts and methods is possible or desirable?

What are the prospects for exchange of data tapes between institutions?

There will be a coffee hour at 4:00 with Department of Botany staff

Digitized by the Hunt Institute for Botanical Documentation

May 5, 1969

INFORMAL CONFERENCE ON DATA PROCESSING

IN THE HERBARIUM

DEPARTMENT OF BOTANY, SMITHSONIAN INSTITUTION

The early formative years of the world's great museums and botanical gardens were devoted to building up plant collections. There are now about 17 institutions in Europe and North American with holdings of over 2,000,000 specimens. Sheer size has long been considered a sign of excellence, but few would argue that further signicant increases in numbers of specimens in large herbaria actually enhance the value of the collections proportionately. On the contrary, size alone brings into sharper focus two immediate problems: the ever-mounting costs of collection management, i.e., the cost of technicians to process and file specimens, loans, and exchanges, as well as the storage costs, herbarium cases, and additional floor space; and, secondly, the realization that greater size does not lead to easier or more complete accessibility of data.

The herbarium is in effect a large data bank, a kind of sheet file arranged phylogenetically and/or alphabetically by taxon, family, genus, and species. Much of the data in herbaria are effectively locked up unless multiple files on locality data, collectors, etc. are concurrently maintained, a near hopeless task. The time is rapidly approaching when taxonomists must look for new means to make data from herbarium collections available to other botanists or plant users, as well as expedite the storage and dissemination of taxonomic research. The anticipated needs of the Flora North America project alone will amply justify any efforts in this direction.

If botany is to follow the development of information retrieval in other fields, then the most important tool will be electronic data processing (EDP). In simplest terms, the computer enables us to construct multiple card files to taxon, geography, collector, date of collections etc., or any combination of these descriptors, which may be printed out on sheets or on cards.

The storage capabilities, versatility, and speed of large, modern computers are unquestioned, and the kind of information recorded for botanical specimens seems tailor-made for computer input. Why, then, have so few herbaria explored data processing as a collection management and retrieval system, and for that matter why is there in operation no computerized data bank for any major herbarium?

Digitized by the Hunt Institute for Botanical Documentation

These problems should be more clearly delineated and, in part, solved within the next two or three years. A group at the University of Colorado has laid much of the groundwork for the concepts and terminology in this new science, and the Smithsonian Institution Information Systems Division has built up considerable experience in source data automation of museum collections. The day will come when institutions desiring to explore computerization will have access to packages of programs tailored for herbarium applications.

Given computers, programmers, and the money to support them, one faces a final problem, that of putting the information on botanical specimens in machine-readable form. In the present stage of technology this means, briefly, punching the data on cards or paper tapes in fixed fields and feeding them into the magnetic tape or disc of the computer. For most herbaria there are important and often crucial problems of computer hardware, its availability and cost. Computers are expensive, and unit cost does not go down as volume of data increases. The programmers form an indispensable link between computer and user, but they, too, are expensive and often in short supply. Computer costs include actual retrieval operations, printout, and the programmers.

The main reason perhaps for the slowness in adopting EDP is that most questions on computer techniques for storage of information on botanical specimens have still to be asked, much less answered. For example, before we decide what to store, we should ideally know what kinds of questions a taxonomist or other user will ask, how often he will ask them, and in what form they should be. If useless or irrelevant data are stored, then an expensive stage of input is wasted. If needed data are not entered, it may be too time consuming to add them.

The purpose of this informal conference on data processing in the herbarium is to bring together a small group of curators who have actually begun EDP studies or who have expressed interest in them. The main theme is communication, definitions, and a mutual understanding of the problems involved, stressing the theoretical aspects rather than nuts and bolts operations.

Digitized by the Hunt Institute for Botanical Documentation

Notes for May 5, 69 Discussion
Smithsonian

Data processing of herbarium holdings.

1. When developing TAXIR asked – what was wanted in herbaria – kinds of data, needs of curators, monographers, floristic workers.
2. Answers seemed generally consistent –
   1. For curator – label data, herbarium identification.
   2. " monographer – above info, plus chars of plant
   3. " floristic workers, – essentially same as monographer, but in less detail.
3. Seemed, then, to indicate building of banks with label data for curators, and some way so that monographers + floristic workers could add on to whatever info was already accumulated by curators.

Note: specific to ...

4. Also needed was complete flexibility of data format, so that different curators could be free to describe their own collections precisely as they wanted, without having to conform to a format imposed from above.

5. Our studies indicate that the important issue was production of a good, common, complete system, capable of sufficient flexibility to fit formats as desired.

6. As indicated by Mason in his advanced notes, the most expensive part of the effort is actually the gathering of the info from the specimens.

Digitized by the Hunt Institute for Botanical Documentation

assessing the most efficient processes. This study includes personnel needs, data formatting, most efficient machinery, cost/effectiveness.

4. For machinery, the most efficient is the key-tape. [Honey 700] $130/month, with maintenance. Speed over nearest other device (key punch) about 2 times, and over paper tape, about 1/3 faster. - Other features are excellent sensing capacity for correction of input errors.

These studies ~~include~~ based on actual work with Bryophyte specimens, over some 3-4,000 actual specimens, with comparisons of key tape, key punch, ~~and flexowriter~~.

Each specimen described with 32 characters. Also, work on the complete data bank of the genus involved, with 27 characters (descriptors) for ea. specimen.

See the printout -

Question - how can we get the big herbaria into computerized form?

We are applying for NSF funds to mount all the collections of the U. Colo museum, in all, about 1.3 million collections in botany, zoology, anthropology/archeology and geology.

Digitized by the Hunt Institute for Botanical Documentation

Some simple handles as low as 7¢/specimen.

22¢ – of which ca 2.5¢ is computer cost.

Cost – about ____ ¢ specimen, for all operations.

Question – how can we meet costs –

1. Must get the directors of museums (organization) to agree on procedures.
2. Get the assistance of the community to aid – Asa Gray person who honors specimens, to provide a predetermined set of curatorial characters –
3. If, as Nasons suggests, there are 17 herbaria in US & Europe with over 2 million collections, there are probably 7-8 in U.S. + Canada and if, as we estimate, ~~these~~ the costs will be $3-400,000 per 1.5 million, then a guess as to overall costs will be ~~10-15 m~~ 3.2 million.

But, this fig. can be lowered, with continued effort + improved techniques.

Cornell, NY, Wash, Chic, U.S., Berkeley.

8
400,000
$3,200,000

Digitized by the Hunt Institute for Botanical Documentation

Stoppa

DATA ON LABELS

GENERAL CONSIDERATIONS

Before a program for the storage and retrieval of data is written, it is important to make it as compatible as possible with the idiosyncrasies in the documentation of past and present herbarium label formats. The chief difficulty in developing an efficient system for electronic data storage is the incompleteness of available records. Many herbarium specimen labels suffer in this respect in one or more items of information. In addition, the various parts of an individual record should not be too long and should not have to be coded before input either by the botanist or by the machine operator. Coding adds further chance for errors and takes time from the input operation. Some types of coding can be handled by the computer, such as the coding of scientific names by means of a "taxonomic dictionary."

Because the working up of the "backlog" of data on available specimen labels will be hindered by imcompleteness and illegibility of the labels, it is important that efforts be made as soon as possible to improve the information which is becoming available on all the new specimens being added to herbaria day by day. Therefore, we have considered in detail the data associated with herbarium labels from the various angles which have a bearing on the development of a computerized information retrieval system. The items discussed below are mostly marked to indicate into which of three categories we believe they fall; some we consider essential (E), others desirable (D) and the rest optional (O).

LABEL FORMAT

The dimensions of the label do not need to be standardized but in the case of bryophytes and lichens stored in packets or envelopes, the label size should not exceed that of the container. Normally there will be no limit to the choice of format for the label. The order in which the information is arranged may be restricted if the labels are produced as a byproduct of recording data on a preprinted catalogue record form. Each herbarium would be free to determine its own label format and recording system while still able to participate in a nation-wide data storage project.

PLANT NAME

(E) The label must bear the scientific name of the taxon with at least the binomial consisting of the generic epithet, specific epithet and author name(s). The name should appear near the top of the label where it may be seen easily.

(D) If the identification has been carried to a lower level than species, it should be so stated by adding the epithets and author name(s) down to the lowest level, using ssp, for subspecies, var. for variety and forma for form. Author name(s) are omitted for infraspecific taxa which represent the typical element of the taxon and repeat the preceding epithet.

(O) Since specimens with imcomplete identification may constitute valuable records in a data system, it is acceptable to give the name of the genus followed by "sp." which signifies "species not determined".

O

Digitized by the Hunt Institute for Botanical Documentation

Orthography. - No blank spaces should occur within an epithet. Special abbreviation such as "cf.", "aff.", "prope" or "?" should not be used between the generic and specific epithets but comments of this type may be placed under "Notes". Abbreviations for authors' names should be consistent with the usage in an up-to-date flora or manual.

Synonymy. - There will always be disagreement about the hierarchical status of some plant names and also there will be in use a certain amount of valid synonymy. It is hoped that the C.B.A. checklists now in preparation will serve as guides to synonymy. If synonyms are to be shown at all on labels, they should be given under "Notes".

Hybrids. - Interspecific hybrids should be described by the single name (if available) without multiplication symbol (X) after the generic epithet or, if necessary, by the scientific names of the two parents separated by X.

Types. - If a specimen constitutes TYPE material, the scientific name should be followed by the statement "TYPE" and more particular information about the kind of TYPE (ISOTYPE, HOLOTYPE, etc.) may be put under "Notes". Definitions for kinds of TYPES may be found in McVaugh et al., 1968.

Numerical Coding. - The coding of a taxon by number is not necessary but, if desired, it may be used in local information systems to develop a more rapid input of specimen data or sight records. It is also useful in a simple system based on tab cards where economy of space precludes the use of the full scientific name.

LOCATION

(E) A locality name (word reference) and a grid location (numerical coordinates) must be given on the label. The locality name should be an accepted geographical name of a settlement or a physiographic feature as listed in one of the official gazetteers. Distance (in kilometres or miles) and direction (compass bearing or cardinal point) from the place named are optional statements. If a local name is not specific (e.g. "Gull Lake") then the township and county should be added. The numerical reference will help pinpoint the location when the local name does not do so.

(E) The grid location must be one of the three indications, namely the latitude and longitude to the nearest minute, the U.T.M. (Universal Transverse Mercator) grid reference to the nearest 10 X 10 km square, or the 1:5,000 map sheet number followed by the U.T.M. reference to the 10 X 10 km unit by means of two digits.

(O) The label may also contain some descriptive regional or local names. The first is often given in the title, e.g. Flora of the Maritimes, Plants of the Bruce Peninsula, etc. The name of the Province or Territory, County or District, National or Provincial Park or title of a physiographic or vegetational region may be of interest in floristic studies.

Altitude. - The altitude may not be of great importance in lowlands, but for mountainous regions it becomes an essential part of the label. Unless a special study is being made, an estimate to the nearest hundred metres is usually sufficient.

Digitized by the Hunt Institute for Botanical Documentation

Cultivated material. - When a specimen has been collected in a field plot, test garden, greenhouse or growth chamber, no location should be given in the space ususally allocated for this information. The statement "Cultivated - see Notes" should be inserted and the details given under "Notes."

HABITAT

(D) Some statement of the nature of the environment in which the specimen was collected is important in many cases. A standardization of this type of information does not seem feasible. A short characterization might stress either the topographical or pedological aspects or a specific factor complex such as the type of human disturbance or the dominant life form of the habitat. Lists of dominant plants or associated species should be relegated to the notes but the name of the host plant for an epiphyte or a parasite should be given as part of the habitat description.

DATE OF COLLECTION

(E) The day, month and year of the collection must be stated and in that particular order.

(O) To simplify input procedures, dates on catalogue cards or on tab cards and paper tapes may be given in arabic numerals separated by slashes, e.g. 5/7/1968. There are sufficient old collections to necessitate the use of more than the last two digits for the year.

COLLECTOR & COLLECTOR'S NUMBER

(E) The label must contain the initials and surname of the principal collector of the specimen. If more than one person is listed, the person whose number is given must be cited first.

(O) Titles such as Mrs., Prof., Col., Dr. etc. should be omitted from labels. If the collector is, for example, a Mrs. M.I. Smith, then the initials must be her own and not those of her husband, otherwise there is no distinction when the title is dropped.

Names of additional collectors should be limited to two and the statement J.H. Smith et al. is of little value on the label itself. It has its proper place in abbreviating a list of three or more names when information from a label is cited in a publication.

(D) Each collection (except for exact duplicates in time and space) should receive a unique number allotted by each collector in chronological sequence. Collectors should not begin again at number 1 each new season. Numbers involving coding for geographical regions (e.g. 529AG for a specimen collected at Arkona Glen) or including dates (e.g. 68529 for number 529 collected in 1968) should be avoided. Collector's numbers should not contain blanks, dashes, slashes or any other symbols. It is acceptable, when necessary to designate divided material resulting from accidental mixed collections, to append a single alphabetic character to the number (e.g. 12529A and 12529B) for two different taxa originally assigned the number 12529).

HERBARIUM CODE & ACCESSION NUMBER

(E) Each herbarium sheet is normally stamped with the distinctive seal or stamp of the herbarium to which it belongs. The internationally accepted codes (e.g. CAN, OK, SASK. TRT, etc.) are ideal for use in data processing and these codes are usually not more than four letters in length.

Digitized by the Hunt Institute for Botanical Documentation

(O) In order to designate a particular sheet in a herbarium, it is necessary that each sheet be stamped with a unique number in a single series for that herbarium. This number is the sheet number or an accession number and bears no relation to collector's numbers. It is not sufficient to locate a specimen by the collector's number because so-called "duplicates" have been known to e a different taxon or to contain a mixture. More important, however, is the fact that many collectors have never used collector's numbers and others have repeated their numbering sequence in successive seasons.

DETERMINATOR

(D) The label should bear the name (initials and surname) of the person who identified the specimen, preceded by "Det." and followed by the year in four digits. Any subsequent change of name should be made on a standard annotation label glued to the sheet near the original label. All such annotations should be made in ink or by means of printed correction slips and should bear the name of the revisor and the date.

NOTES

(D) It is useful to note near the bottom of the label the number of replicas prepared from the same collection. The label may also allow a space of two or three lines at the bottom, possibly with a printed heading "Notes", where additional information can be given. Statements about the height of the plant, colour of flowers, abundance, local names or uses, references to photographs, chromosome counts or samples of seeds taken can be recorded here. In a data bank it may not be feasible to search this information for specific features, but the Notes can be stored and recalled in toto.

The name and number of an exsiccata specimen should be entered under notes and the title of an expedition may also be useful information.

In the case of TYPE material, bibliographic reference for the new taxon should give the place and date of publication for ready reference.

-- Extract from report by R.E.Beschel & J.H.Soper,
March 31, 1969, Committee on Automation and
Standardization of Herbarium Procedures,
Canadian Botanical Association.

Digitized by the Hunt Institute for Botanical Documentation

Hale + Creighton

May 1, 1969

Summary of The
Botanical Type Collection and Dissemination
System (BOTCAD)

Purpose of the System

An automated information collection and dissemination system is being developed to create a data bank containing "type" data derived from botanical specimens throughout the world. A record for each "type specimen" is prepared that contains:

1) historical facts concerning the specimen and

2) a list showing the institution acronym, catalog number, and kind of type reported by each herbarium for duplicates of the specimen.

Each record, printed on a 5 by 8 index card, is also made available on magnetic tape. A replication of the data bank will be furnished to each institution that participates in the project.

Methods

Data, derived from botanical types at the NMNH are entered into the system on a continuing basis. Records representing selected genera are printed on index cards which are mailed to participating institutions. A two card data record, perforated for easy burst, is furnished for each species. The first card is a "Record of Holdings" which the recipient may wish to keep on file until it is superseded by a more comprehensive record. On the first distribution, the cards may contain only data from the NMNH. Subsequent distributions will reflect the holdings of all participating institutions that have reported a kind of type for that taxon.

The second card is a "Data Collection Form". If the recipient has in his herbarium a duplicate of the specimen cited on the first card, he is asked to complete line 12 of the second card (Institution acronym,

Digitized by the Hunt Institute for Botanical Documentation

catalog number, and kind of type), and to detach and return the card. If the recipient has other verified types, especially of taxa recently described, he may enter the data by completing lines 1 through 12.

"Data Collection Forms" returned to the Smithsonian will cause the system to update any type already on record and to incorporate any reported type that is new to the file. All data thus reported is reflected in subsequent distributions for the taxon.

Management Information Reports

The system automatically adjusts its activities according to the status of records in its data bank and provides the following management information:

1) Genera activated for a dissemination-collection cycle.

2) Genera fully updated and disseminated and consequently deactivated.

3) Institutions that have been furnished current systems data for selected genera and asked to respond.

4) System's holdings by family, genus, etc.

5) Each institution's holdings by family, genus, etc.

6) Other miscellaneous items of information.

The system prevents redundant mailing of records (unless otherwise directed) and insures that no institution in the network is by-passed regarding any group, when the various herbaria report randomly on genera of their choice at differing points in time. The system rapidly processes large volumes of data; facilitates data validation and correction; and prints accurate and timely reports on the collective holdings of herbaria throughout the world.

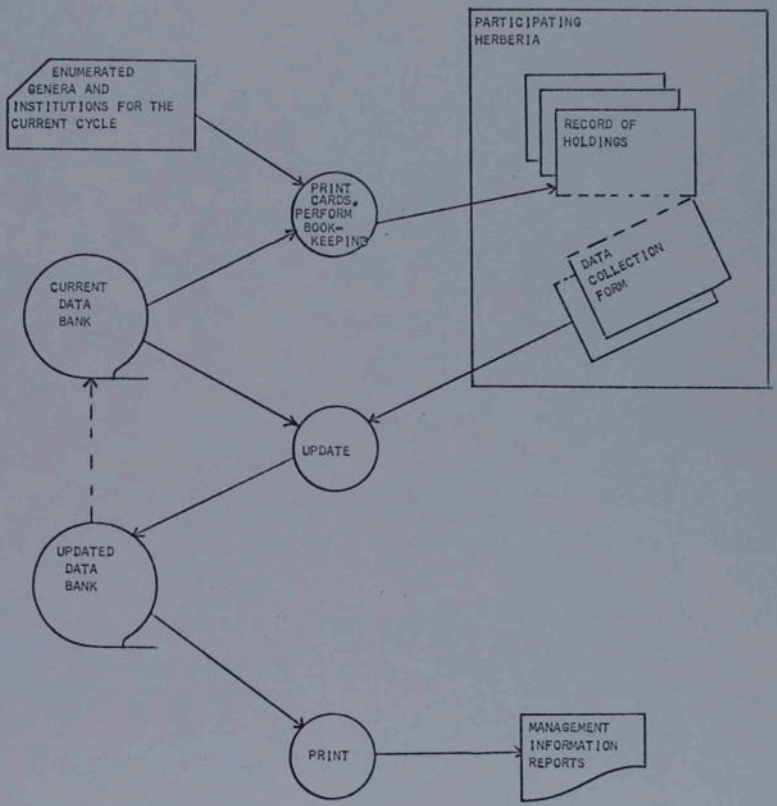
The stringent housekeeping and bookkeeping chores involved in a comprehensive information collection and dissemination system are in large measure handled by the computer.

Digitized by the Hunt Institute for Botanical Documentation

April 30, 1969, ISD

Botanical Types

Schematic of Data Collection and Dissemination



Digitized by the Hunt Institute for Botanical Documentation

TYPE REGISTER - CATALOG CARD

SCROPHULARIACEAE

MIMULUS DIFFUSUS GRANT

ANN. MO. BOT. GARD. 11:254. 1924.

COLLECTORS: JEPSON, W. L. HALL, H. M. COLL. NO.: 1959

DATE COLL.: 29 MAY 1901

US: CALIFORNIA: SAN DIEGO CO.: PALOMAR

US...0412781 ISOTYPE

THIS CARD MAY BE RETAINED FOR YOUR FILES

RECORD ID. 313

TYPE REGISTER - DATA COLLECTION CARD. 38552030821787249024

IF YOU HAVE IN YOUR HERBARIUM A DUPLICATE OF THE ABOVE SPECIMEN, COMPLETE LINE 12 BELOW ON THIS CARD.

IF YOU HAVE OTHER VERIFIED TYPES, ESPECIALLY OF TAXA WHICH YOU HAVE RECENTLY DESCRIBED, PLEASE ENTER DATA AS COMPLETELY AS POSSIBLE ON THIS OR OTHER CARDS AS NEEDED. RETURN TO DEPARTMENT OF BOTANY, SMITHSONIAN INSTITUTION, WASHINGTON, D.C. 20560.

01. FAMILY: RECORD ID. 000313
02. GENUS: SPECIES:

03. VAR. OR SUBSP.: FORMA:

04. AUTHOR:
05. CITATION:

06. COLLECTORS: COLL. NO.:
COLL. DATE:

07. LOCALITY: COUNTRY:
08. STATE OR PROVINCE:
09. COUNTY OR DEPARTMENT:
10. TOWN OR SUBREFERENCE:

11. REMARKS:

12. INSTITUTION ACRONYM: CAT. NO: KIND OF TYPE:

Digitized by the Hunt Institute for Botanical Documentation

Crovello

Table 2. Problems in the use of EDP in biological collections.

A. Recognition of those uses of EDP that have scientific value.

B. Physical
   1. Existence of necessary equipment (hardware and software)
   2. Availability of equipment to curators
   3. Compatibility of an EDP system with different systems at other collections
   4. Availability of space (for equipment and for data)

C. Cost
   1. Price of equipment
   2. Price of trained EDP personnel
   3. Source of funds for EDP
   4. Price of time demanded of the curator to initiate and maintain an error-free system of information

D. Human
   1. Fear of EDP by curators
   2. Ignorance of EDP by curators
   3. Inertia generated by awe at magnitude of certain EDP tasks

E. The collection
   1. Adequacy of the collection to produce reliable results
   2. Specimen misidentification
   3. Synonymy
   4. Choice of the particular specimens to process

F. Organizational level of institutions for studies using EDP

G. Optimal level of taxonomists organization for studies using EDP

H. Generation of a suitable context to decide if use of EDP is worth the expenditure of the required resources in a particular situation

I. Codification of biological names

Digitized by the Hunt Institute for Botanical Documentation

Crowell

April 24, 1969

Table 1R. Revised Field Format for Encoding Data from The Edward Lee Greene Herbarium.

Field Number	Information In Each Field
1	Post 1966 accession number
2	Herbarium Greeneanum number
3	Genus number (from Dalle Torre and Harms)
4	Binomial (or trinomial) and author.
	Amount of each plant part present
5	Leaves
6	Flowers
7	Fruits
8	Seeds
	1 = absent 2 = immature 3 = at least some mature
9	Type specimen and who designated it first? 1 = not indicated as such 2 = holotype 3 = isotype 4 = lectotype (a second digit is added to type specimens to indicate who first designated it as such; e.g., 21 means annotator one first designated it as a holotype.)
10	Collector's name and his collection number
11	Date (mo/day/yr.)
12	State (standard abbreviation)
13	County (standard abbreviation)
14	Locality and Habitat
15	Determined by?
16	Comments (as explained on page 14c of the proposal)
17	Annotations by first annotator (either the author, or a subsequent worker) if any.
18	As field 17, but for annotator 2, if any, etc.

Digitized by the Hunt Institute for Botanical Documentation

Crovello
13 May 1967

Table 1. Uses of EDP in biological collections according to the source of information.

Label Information

Catalogues, or cross reference lists (e.g., by taxon, location, date, etc.).
Collection inventory.
What we have, when and where we got it, what accessioned material cannot be located, what material is on loan to whom, what taxa are not represented in the herbarium and should be, what collection numbers of E. L. Greene do we have, what types do we have, etc.
Lists of specimens available for exchange or loan.
Lists of specimens examined or scored in a taxonomic study.
Preparation of duplicate ribbon copies of specimen labels.
Breeding and flowering times.
Geographic distribution maps.
Multivariate analysis of distribution maps.

Specimen Information

Automated character recording.
Storage and analysis of information for floras.
Preliminary establishment of multivariate patterns of phenetic relationships among groups.
Automated construction of keys using generalized discriminant functions (Blackith, 1965).
Automated specimen identification.
Screening and searching data for special characters or particular states.
Host-pathogen lists.
Ecological studies, e.g., principal components analysis of the vegetation of a small area.

Extraspecimen Information

Nomenclatural lists, e.g., the Gray Herbarium Index.
Bibliographies according to taxa, geographic areas, etc.
List of current taxonomic projects, by taxon, taxonomist, country, etc.

Digitized by the Hunt Institute for Botanical Documentation