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The Hunt Institute Archives contains hundreds of thousands of pages of historical content, writing and images, created by thousands of individuals connected to the botanical sciences.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of time and social context in which these materials were created, some of the collections contain material that reflect outdated, biased, offensive and possibly violent views, opinions and actions.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 Archival records are historical documents, and the Hunt Institute keeps such records unaltered to maintain their integrity and to foster accountability for the actions and views of the collections'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.

UNIVERSITY OF CALIFORNIA, IRVINE

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

DIVISION OF BIOLOGICAL SCIENCES
DEPARTMENT OF POPULATION AND
ENVIRONMENTAL BIOLOGY
IRVINE, CALIFORNIA 92664

November 8, 1968

Dr. D. J. Rogers
Department of Biology
University of Colorado
Armory 101
Boulder, Colorado 80302

Dear Dave:

Thank you very much for your letter of October 29th and the enclosure. I was most interested to receive a copy of your paper on the uses of computers in taxonomy, I had been grossly ignorant of the existence of this paper. It is one thing to have all these ideas, but it seems to be quite another matter to get some of the more staid members of certain of our more distinguished institutes to consider them.

Regarding the Index Kewensis, I have expressed myself badly. I have worked for a number of weeks at a time in both the Gray Herbarium and the New York Botanical Garden and their arrangement isn't the same, or wasn't the same when I saw it, as at Kew. If you remember the Kew situation, the Index is spread right across the whole of a very long rostrum, or at least it was until they moved to the new wing this year. However, it doesn't really matter because judging from what was going on when I was over at Kew this last August, it's going to be very difficult to find anyone willing to succeed Bullock in editing the Index.

Things are more frantic than usual here, I couldn't get away for your taxometrics meeting in Boulder. It was extremely interesting having the privilege of seeing your Taxir program last month, and I hope we may have the opportunity of returning the compliment in the not too distant future. Meanwhile we are struggling to get our program back on our system. We unfortunately are confronted with a large number of teaching terminals and the time sharing system for them, and while we can get batch processing done, it is a little difficult to get time for on-line programs.

Yours sincerely,

Arthur S. Boughey, Chairman

ASB:ss

29 October 1968

Dr. Arthur S. Boughey
Population & Environmental Biology
University of California
Irvine, Calif. 92664

Dear Arthur:

I finally and belatedly have read your manuscript, "Computers in Taxonomy." I appreciate your sending it. Certainly the concepts given there do parallel my feelings. I have only two comments that you might care to examine. On page 4 of the manuscript you have indicated that Kew is the only organization which cuts up and files the Kew Herbary under the appropriate genera. As a matter of fact the New York Botanical Garden and the Gray Herbarium both do this. I believe as a matter of fact the idea originated in New York with the former curator, E.D Merrill.

Another comment is that you might be interested to read our paper,, "Uses of Computers in Studies of Taxonomy and Evolution," a copy of which is enclosed.

Other than these comments I find it an interesting summary of activities but one which I believe could be expanded today.

We look forward to seeing you next week.

Sincerely,

David J. Rogers
Professor of Biology

DJR:gm

UNIVERSITY OF CALIFORNIA, IRVINE

BERKELEY • DAVIS • IRVINE • LOS ANGELES • RIVERSIDE • SAN DIEGO • SAN FRANCISCO

SANTA BARBARA • SANTA CRUZ

DIVISION OF BIOLOGICAL SCIENCES
DEPARTMENT OF POPULATION AND
ENVIRONMENTAL BIOLOGY
IRVINE, CALIFORNIA 92664

October 10, 1968

Dr. D. J. Rogers
University of Colorado
Taximetrics Laboratory
Department of Biology
Armory 101
Boulder, Colorado 80302

Dear Dave:

Thank you very much indeed for your hospitality at Boulder and for providing me and Kent Bridges with the opportunity of seeing 'Taxir.' Enclosed is the draft of something I put together over a year ago which I may or may not attempt to get published. As I mentioned at the time, I wanted you to see this because some of our ideas were bound to overlap.

In the identification program we ourselves worked out, we had of course many of the same problems that you have had in developing Taxir. We went through the verification stage and produced an audit program before we removed the necessity of this altogether by utilizing a character display device. From the start we avoided the restrictions of core storage, such as one has for example with the very simple little program which Goodall published, using peripheral storage devices.

Although identification was the prime purpose of our program, we have always intended to develop it into many other operations, including for example what you have called your 'Book' operation. In doing this we were, however, handling species rather than specimens.

We were most interested to learn of some of your special program developments, for example, your compression device, and it was especially interesting to hear the various cost estimates, a thing we have never looked into ourselves.

May I wish you all success in the development of your very fine program and the pursuance of version 2.

Best wishes.

Yours sincerely,

Arthur S. Boughey, Chairman

ASB:ss
Enclosure

COMPUTERS AND TAXONOMY

A. S. Boughey

University of California, Irvine

Taxonomic principles and practice are currently undergoing a critical and exhaustive re-examination and revaluation. It is no longer a question simply of recognizing the de facto existence of several levels of categorization, Turrill's alpha to omega taxonomies (1951). The discovery of the method of genetical coding in living organisms, and of the mechanisms for the transmission of this genetic information, has challenged many current assumptions as to the biological nature of species. As Ehrlich (1964) has pointed out, one consequence of this genetical work has been to remove the classical distinction between genotype and phenotype. Another has been to invalidate the universality of the biological species concept. Certain traditional goals, such as the cataloging of organic diversity, now appear ill-defined and perhaps even to some scientists ill-advised.

Simultaneous with this critical re-evaluation of taxonomy, and contributing to it in no small measure, has come the development of data processing with electronic computers. The use of computer techniques, and the taxonomic implications of such applications, are sometimes treated simultaneously. Thus the publication by Sokal and Sneath (1963) of a treatise on numerical taxonomy, followed more recently (Sokal and Sneath 1966) by a comprehensive statement of the employment of data processing to taxonomic procedures, evoked considerable reaction among traditional taxonomists, most of it unfavorable, some strongly antagonistic. Such reactions as that published by Kalkman (1966), although understandable, tend to obscure the great opportunity which the application of data processing techniques offers to taxonomy, even without involving the replacement of the present system by a numerical one, or requiring of taxonomists

a deep mathematical knowledge and wide computer experience. The issue of phenetic versus phylogenetic classifications will ultimately be resolved, but this is no reason meanwhile for protaganists of traditional views to ignore the potential utility of the main instrument of their opponents. Williams and Lambert (1965) have remarked that the rationale of numerical techniques for the classification of elements into sets has been widely misunderstood. They point out that this misunderstanding has not only evoked the strictures of taxonomists (Seal, 1964), but has produced an exaggerated respect for classification based on numerical taxonomy.

Particular uses of electronic data processing (EDP) in biological collections have recently been very completely and competently reviewed by T.J. Crovello (1967). This comprehensive analytical review of the uses of EDP in taxonomy is supplemented by cost data which are very relevant to any question of contemporary EDP applications. It has to be remembered however that we are now on using third generation computers, and that each succeeding generation has seen a reduction in cost of computational time to approximately one-tenth of that of the preceeding generation. It is the saving in cost as well as in time which Kalkman (1966) appears to overlook.

As Crovello describes, computers in a commercial form are scarcely 15 years old. The great majority of todays working taxonomists were trained prior to their introduction, and a large number of major herbaria still lack ready access to computer facilities. Nor is there as yet a new generation of taxonomists schooled in the use of computer data processing techniques as matter-of-factedly as at one time a knowledge of Latin or Greek and of modern European languages was acquired. Workers in herbaria and museums not closely associated with a university or major research institute will in the future be greatly handicapped in their attempts to acquire such a familiarity with computer

techniques, for there will be no computer unit with whose hard-ware, soft-ware and personnel they can readily interact.

Perhaps because of the initial research applications of computers, their most familiar present use in taxonomy has been in multivariate analysis. Papers such as those of Williams and Lambert (1961) Petter (1965), and Jolicoeur (1959) illustrate the use of such techniques. This computer application involves no radical change in taxonomical thinking or approaches. The procedures followed are such as a taxonomist would in any case employ, but in which he can utilize a larger data base, and extract from this more information than would be practicably possible without the use of a computer.

A more revolutionary field, one into which, as Crovello describes, computer uses in taxonomy have scarcely begun to penetrate, is in information retrieval. Crovello describes potential applications under the three headings of label information, specimen information and extra specimen information. In order to explain how information retrieval techniques may be applied to taxonomy, it is perhaps necessary in addition to flow chart the processes which a taxonomist follows when undertaking the examination or revision of a particular taxonomic group or the floristics of a given geographical area. This is applying to taxonomic studies the procedures of systems analysis, which were made immensely profitable with the development of EDP. Classical taxonomists may resent the subjection of their traditional procedures to scrutiny by the techniques of systems analysis. Nevertheless, pressurized as they are by routine demands, even they must welcome the introduction of processes which promise some measure of relief.

Such a flow chart is illustrated in fig. 1. The first stage of the operation is the selection for study of a group or groups, or of a geographical region. Even if the taxonomist is concerned with a region rather than a group, he must ultimately define his area in terms of systematic groups in order to obtain the necessary reference collections for botanical determinations.

Once the groups have been selected, it is usual to embark simultanously on two operations, the assembly of the relevant published literature for the area to be considered, and a listing of the basonyms associated with the group. The taxonomist then continues with these dual operations, locating the types of the various basonyms, while simultaneously extracting from the literature records of the cited specimens of the groups with which he is concerned from the selected geographical area.

In these initial traditional taxonomical operations as is well known to systematists four aids are available. The Index Kewensis and its supplements provide a list of all basonyms published until within the last few years. Kew Hebarium has a sufficiently generous supply of the Index and of space to be able to cut up this work and arrange the cut pages under appropriate genera. No other hebarium is apparently sufficiently well endowed as to permit this arrangement, and various degrees of awkwardness are involved in searching for basonyms through the various volumes of the Index. For North America, the Index Kewensis is supplemented by the Gray Card Index, which is available in most major North American hebaria. Thirdly, the Index Herbariorum lists in many cases the location of major type collections for particular groups, although this is not indexed. Lastly, many major hebaria have copies of a mimeographed document produced some years ago by J. Hutchinson indicating the location of the types of many groups.

Idea originated in NY, and is also done in Gray

Having thus located as far as seems possible the types of basonyms and the deposition of cited specimens, the traditional taxonomist now usually proceeds to write to the various institutions in which these are deposited, seeking the loan of the specimen for examination. For those specimens which cannot be borrowed he arranges a visit, as would have to be done for example in the case of any Linnean type specimen in the Linnean Society of London's collections. Or else he has a friend or colleague in the institution involved examine

the specimen for him, has it photographed, or merely decides he will just have to ignore it. Information on cited specimens is obtained in the same way.

After an examination of the various types and cited specimens, the taxonomist decides either to accept or to revise the previously existing taxa, to prepare diagnoses and descriptions of any new taxa he decides to create, and possibly amplifies descriptions of others.

Having done this, a working key is usually evolved, if only for the use of the individual taxonomist himself, and this is quite commonly eventually worked up into a published key. In any case, this key is usually applied to all the cited specimens which have been assembled before these are returned to the institution from which they were borrowed.

Finally, publications are prepared containing any new taxa or revisions, and listing the distribution of the specimens which have been examined and determined.

This operation may vary somewhat from one worker to another, but the flow chart shown in fig. 1 represents what must take place sooner or later in any such traditional taxonomic study. Apart from the four aids which were quoted above, the memory of the taxonomist, and of his colleagues with whom he discusses the matter, are the sole sources of information retrieval in all these traditional operations. It is instructive to take the flow chart illustrating these procedures and add presently existing electronic data processing systems which could assist these several taxonomic operations. Fig. 2 illustrates diagrammatically the various stages at which particular processes can supplement the memory and minimize the time expended by the taxonomist.

The first computer operation concerns access to the literature of particular groups or regions. The IBM Company developed some years ago a program known as KWIC, which are the initials for 'Key Work Information Coding', a little known and not yet extensively employed method of arranging literature references

(Luhin, 1959). Once data files have been assembled by the KWIC process, information as to literature on a genus, family, or an area is very rapidly located. Fig. 3 illustrates a section from the kind of data file which could be assembled by library personnel with no specialist knowledge of taxonomic processes, but having full access to taxonomic literature. A file such as this could be produced in any major hebarium fortunate enough to possess representative holdings of the world's taxonomic literature. The data files so assembled could be made available for those less fortunate, which would greatly facilitate the compilation of information by isolated workers and individuals in smaller institutions.

The KWIC Index can similarly be used to compile a list of basonyms from the Index Kewensis, the Gray Card Index, and from current literature. A print-out of a given set of basonyms could then very quickly be automatically extracted. The availability of such a basonym file might possibly have a minor effect on eventual sales of the first two indices, but is unlikely to have any short-term effect. A KWIC Index of basonyms could be revised very simply and at little cost on say a three-monthly, six-monthly, annual, or even a continuous basis, and would therefore be very much more up to date than it is possible to keep the indices presently published. It could also be made available to anyone with access to a computer of reasonable capacity. Alternatively, as Crovello suggests in regard to the Gray Index, the print-out can be mailed out from the central herbarium at a small charge.

A very recently developed program devised by D. Gilbertson (private comm.) offers a very easy and extremely quick method for editing a basonym file. The required file entry is selected and reproduced on a CRT display unit, in this instance an IBM 2250. Any necessary deletions and additions are then made to the entry, and checked visually against a duplicate original entry before returning the item to the file. If this file is maintained on disc or tape, the

cumbersome routine of manual card shuffling, repunching, and verification is totally avoided.

The location of types of basonyms, where known, can be linked to a KWIC Index of basonyms. The preparation of this addition would most desirably be carried out on an international basis, and would require personnel rather more taxonomically knowledgable than those who could prepare the first two KWIC Indices discussed. Once constructed however, the file could be regularly updated as publications appear, and would save incredible amounts of time for professional taxonomists.

The next stage in the completion of a taxonomic exercise, which involves the reference to diagnoses and original descriptions of basonyms, is one in which data processing techniques are most useful, and already highly evolved as Crovello describes. While again this would require an international effort to complete the data file in a reasonable time, once this has been compiled, the data can be stored on tape or disc and made available to any institution. When this has been done, in order to obtain the reference for original descriptions, instead of having to go through the very laborious process of referring to source volumes, or visiting areas where such works can be inspected, references could be obtained in seconds. In the absence of a linked microfilm system, copies of the articles referred to could then be obtained from one of the institutions fortunate enough to possess this literature. A 'teleprinter' type of machine is now in experimental use in the U. S. which transmits a copy of a document to the 'print-out' machine over a telephone link. The provision of such facilities would make the process of literature search a matter of seconds for any individual with access to a computer console, and the obtaining of a photocopy of any particular item a matter of minutes if the necessary 'teleprinter' or microfilm file is available.

Still restricting the discussion to existing computer programs, the penultimate stage in the flow chart diagram is the preparation of distribution maps. This process has been computerized for some years, and follows a program which for example has been used by the British Ecological Society to indicate the area of distribution of British plants (Perring, 1963).

Finally, a list of specimens with accompanying diagnoses, descriptions, illustrations and maps can be entered on tape, from where it may be retrieved in a matter of seconds at any time when required for any further revisory work.

Turning to those aspects of computerization where only experimental or hypothetical techniques are yet available, future developments become more controversial, and inevitably involve the larger question of numerical taxonomy, upon which it is not intended to embark in this present article. Such matters were in any case admirably and imaginatively presented by Sokal and Sneath (1966) and touched upon by Crovello (1967). It will be obvious that once specimens have been assembled, data processing techniques employing some method of numerical taxonomic analysis could sort these into phenons, thus establishing entirely new taxa in some cases, or possibly confirming previously recognized taxa in others. In the same numerical taxonomy processes, keys can be prepared for these phenons, and the processes for the identification of phenons can be made more extensively automatic and capable of repetitive confirmation. All these aspects were considered by Sokal, Sneath and Crovello in their works previously cited.

In this institution a less controversial experimental program has been developed for the identification of conventional taxa using a multi-access key system, and this will be referred to later. It might now be appropriate to examine some of the current taxonomic projects which are proceeding in various parts of the world and the ways in which storage and retrieval by computers could greatly assist and speed up these various activities. For example, an AETFAT Bulletin is produced each year in Brussells under the direction of

Dr. J. Leonard (1965), which lists all new taxa described from Africa in the proceeding year. Without having been associated with this work, it is only possible to imagine the labor involved in this extremely valuable exercise. The use of computer storage and retrieval in the compilation of such indices greatly reduces the labor involved, and actually increases the accuracy of the work. Literature is surveyed as it is received, new taxa being entered routinely on the storage file. These can be coded for the binominal, for higher classificatory and for geographical categories, thus facilitating retrieval by KWIC and other programs. When a new list is required, a search can be instigated systematically, category by category, and a program can readily be designed to provide a print-out of any desired list from the storage files.

However, much more information than this could be retrieved from such a data file. For example, in the preparation of a flora of a given region, a record of all the basonyms involved, and cited specimens, with references to them, could be prepared from the master file. The compilation of a check-list for a given region, the location of all cited specimens from that region, the obtaining of a distribution map for any particular taxon, all become automatic operations completed in seconds or less.

Various other retrieval techniques are being devised by investigators concerned with the development of information retrieval procedures. A feasability project at this institution (Boughey, 1968) has explored the possibility of a sophisticated search process potentially of very general application. This program provides in the first instance a multi-access computer identification scheme, as illustrated by the flow chart in fig. 3. The first stage in this scheme involves a suitable arrangement of the information to be entered in the data file. This information had to be translated into a language simulating as far as possible natural language. In developing the data file and the language, the remarkable deficiencies of traditional taxonomic descriptions were under-

scored. Terms such as 'round', 'circular', 'orbicular', 'spherical', 'rotund', were found to be used indiscriminantly, apparently according to the particular preference of the individual investigator. Some of these synonyms were taken care of by inserting an equivalence table into the language, which automatically equates designated taxonomic terms. Even then many features of conventional entity descriptions had frequently to be omitted from the computer search because of vagueness and inconsistency, and merely entered so as to be extracted with the final print-out.

The fullest use of this computer identification scheme can only be obtained with the employment of a CRT display unit equipped with a light pencil. The preparation of a data file then becomes a very rapid process. Description categories are thrown on the screen, the taxonomist selects a suitable state specification with the light pencil, enters a value in the same way, and passes on to the next description. Entries can be checked at any time by requesting a print-out, and additions or alterations made. Identification of particular taxa proceed by a reverse process, which it is hoped greatly to speed by incorporation of 'optimization' functions in the program. Such functions select the descriptive categories best examined as the search proceeds, that is, those which will most rapidly route the search to its desired end.

The strategy which has been developed in this search process can be applied to any biological group and offers great promise for both the sophisticated and non-sophisticated user. The multi-access nature of the system enables determinations to be effected on partially complete material. The search process does not depend, as all printed divisory keys do, on the correct selection of one particular character before the user moves down to the next level of the key. Any state specification for which no value can be entered, can be skipped without prejudice to the ultimate answer.

This same search technique offers great promise also in quite a different direction. Most herbaria at some time or another have attempted to record accession data. In the larger organizations it has eventually proved impossible to maintain accession information. If, however, these records were to be stored in a data file, as proposed by Crovello, they could rapidly be retrieved using the same search strategy. Thus, for example, it is a common practice in some institutions to record latitude and longitude of specimens accessioned. The retrieval of all specimens which have been collected at a given latitude and longitude from a large herbarium such as that of Harvard University, the British Museum, or Paris, would be a herculean task of such proportions as could not presently even be contemplated. The storage of the accession information on several million specimens would be well within the capacity of the powerful computers now available, and such a search could be effected at least within minutes, perhaps within seconds. By the same token, it would be possible to produce for example a list of all the herbaceous legumes growing on serpentine soils between one and two thousand meters altitude north of the Tropic of Cancer, or any combination of accessioned information which has been entered in the data file.

While ecological and morphological information is most obviously suitable for such search processes, there would indeed be many further applications of this process which would occur to the practicing taxonomist, and which would develop as specimen notation interacted with computerized techniques. Some further information can be filed by observations on the specimens themselves. Is there, for example, a relationship between meteorological conditions and pubescence, do annuals occur more frequently in the tropics than in temperate regions, is flower color associated with altitude? Perhaps these should be regarded as ecological questions, but many more taxonomical issues can be raised, such as

are opposite leaved plants sympetalous, or what is the relationship if any between wind pollination and leaf hairiness?

In a perfect world in order to identify taxa it should only be necessary to take the Latin diagnoses, which are intended to define the manner in which new taxa differ from previously described ones. Possibly this is indeed a practicable operation in some groups which have been critically treated by the same worker, but it would not be generally possible, because the inclusion also of much of the description is usually necessary for diagnostic purposes. The search program described here can be modified to determine the precise information which must be included in order to provide a unique diagnosis.

Interest in the mapping of species distribution data is increasing. The success of the computer mapping process adopted for the Biological Flora of the British Isles illustrates the possibilities of this technique. To automate mapping on a more general scale would, however, demand some closer regimentation of information recorded on accession labels. Despite all attempts at persuasion or coertion, field collectors are notoriously lax in recording data on herbarium labels. The knowledge that information retrieval processes are to be applied to this material may very well stimulate collectors to a greater and more consistent effort.

Presently there are many floras in preparation in the less developed countries of the world. In the more developed areas, regional flora such as the Flora Europea and the projected Flora North America have been or are being organized. Many processes in the preparation of floras at these several levels, as Crovello has already demonstrated, could be computerized at great advantage. How many times, for example, is the same type material examined in the course of a year by different taxonomists? While this repeated examination is often from a slightly different viewpoint, many of the characters of the type specimens

are unequivocal, and could be recorded in data files, along with illustrations reproducible on display units. It may perhaps at this immediate moment be somewhat too early to claim that photographs of sufficient quality can be stored; as Sokal and Sneath (1966) remark, this is for the future. One can, however, reflect on the number of hours which are spent by taxonomists visiting an institution in order to examine types which curators are unable to permit to be mailed on loan. How valuable would be a compilation of basonyms of a particular area or a particular group? How useful would it be to have a list of synonyms which could be searched and printed-out for given groups in a few moments? How great a help would it be to have original descriptions of species in storage, linked with search processes which can extract almost immediately much information on any taxa which are not readily separable from the material being worked? These facilities, many of them immediately available, could be provided for any herbarium which can be linked by telephone line to a reasonably powerful computer. To these institutions with access by co-axial cable to a powerful computer, the resources available are simply enormous.

The immediate assistance which information retrieval techniques might provide to a regional project such as the projected Flora North America, may be deduced from a consideration of the flow chart in fig. 1. Some but not all potential applications have already been considered by Crovello. Such a regional project differs only in size and scope, not in principle from the taxonomic procedures flow-charted in this diagram. A KWIC Index can be compiled for all basonyms cited in literature from the area, the literature itself being similarly stored and referenced by author and area as well as taxa involved. All known locations of basonyms, and the recorded disposition of cited specimens, can be centrally indexed. The information stored can be made additionally useful if curators supply details of basonym type material in their possession when this is not so recorded in the literature.

Leaving aside the question of whether curators might be willing to have transmitted on demand for display and photography the images of particular type material, presumably they would have no objection to the transmission of particular papers from reference material in their possession. No state in the U. S. lacks a computer center capable of handling these information retrieval and distribution processes, most large cities of any size have at least one. Without moving out of his home institute, home town, or at worst home state, the individual who has chosen or been chosen to work on a particular group for the new Flora North America could expect within several days to narrow down the complete list of basonym and cited material he must first examine before deciding what futher needs to be done with a particular group. He could then construct a data file, and using the search technique already described, determine whether the undescribed material he secures also falls within the described parameters of established taxa. He can correct misidentifications, and create new taxa wherever this seems necessary. He can prepare a well-ordered description following the 'descriptor' and 'state-specification' schedules employed for data file assembly, adding to this the necessary key literature references and distribution maps. The whole description is readily and almost instantaneously retrievable at any stage, and can be edited very simply, rapidly, and at will. Finally, without any effort on the part of the individual worker, his own particular file can be fed into a master file used for monitoring progress on the whole project, or maintaining contact with other contributors working on too large an area, or too extensive a group to be handled individually.