*An offprint from*

The Systematics Association
Special Volume No. 48

# Designs for a Global Plant Species Information System

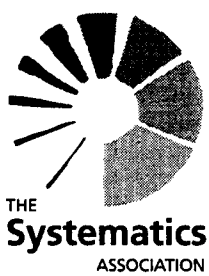
*Edited by*

F.A. BISBY
*Biology Department, University of Southampton, Southampton, UK*

G.F. RUSSELL
*Department of Botany, Smithsonian Institution, Washington DC, USA.*

*and*

R.J. PANKHURST
*Royal Botanic Garden, Edinburgh, Scotland, UK*




Published for the SYSTEMATICS ASSOCIATION by
CLARENDON PRESS · OXFORD
1993

# 9. A view of the future for floristic research

ARTURO GÓMEZ-POMPA and
ORLAY EDWARD PLUMMER
*Department of Botany/Plant Sciences, University of California at Riverside, CA 92521, USA*

**Abstract**

Floristic research faces the challenge of major losses of its study material: plant diversity, especially in the tropics. One outlook on this challenge is to increase our efforts in describing the plants before they become extinct. Another is to modify the forces that are threatening this biodiversity so we can have more time to do the urgently needed research and with luck more plants will escape extinction. The one outlook that we all tend to reject is to continue as we have been doing and let what will be become reality. Accepting this challenge entails a close analysis of the current manner and means of floristic research in terms of how well our current procedures are meeting it and how can we change them in order to better meet it.

## The experience of the Flora of Veracruz

The experience of the flora for the state of Veracruz illustrates the issues. In 1967 the *Flora de Veracruz* was begun (Gómez-Pompa and Nevling 1988). The project faced real and great difficulties because of the scarce resources in Mexico for this type of research—a very poor herbarium and library and a very limited number of trained taxonomists.

But the alarming trends of deforestation and ecosystem conversion occurring in that part of Mexico gave no time to waste with worry. The great demand for ecological and botanical information also required action. But from the beginning, there was a keen awareness of the limitations of classical flora projects in responding to these needs. A flora project is a long-term commitment that goes beyond any one individual's life span.

© Systematics Association, Special Volume No. 48, '*Designs for a Global Plant Species Information System*', edited by F.A. Bisby, G.F. Russell, and R.J. Pankhurst, 1993, pp. 83–93. Oxford University Press, Oxford.

In the 20 years of the Flora of Veracruz we have been able to publish 62 families of the 243 known to exist in the state—approximately 25 per cent of the flora. At this pace the flora will be completed, except for revision, 60 years from now. But the need for the complete information to help preserve and manage these resources is urgent now, not well in the midst of the twenty-first century. This status of the flora is not an exception: the outstanding *Flora Malesiana* for example (Wirawan 1989), has been able to describe only 20 per cent of the expected 25 000 species in 40 years (160 years more to go to completion). A different type of a flora is in order: a flora that would provide information when needed, a flora that would stimulate ecological research, a flora with a built-in training programme for local taxonomists. With these thoughts in mind, we set out to experiment with the format of floras.

*1. Importance of databases*

The first step in the development of a new type of a flora was the use of floristic and environmental computer databases. We knew that they provided the only means for easily updating information and making it highly accessible. But certain issues arising from the use of computer databases had to be confronted.

*(a) Standardization* One was the standardization of the information on herbarium labels that we were inputting. Standardization makes searches easier. The user does not have to think of the several ways in which the information he or she wishes to find can be written—tropical dry forest, selva baja perennifolia, etc. Yet standardization requires interpretation of the available information.

Something that someone might find useful may be lost. A very early attempt regarding this issue was presented in a 1967 meeting organized by UNAM and the Smithsonian Institution (Gómez-Pompa and Squires 1969). However, though great emphasis was placed in that meeting on the need for some standardization, at least of plant names, no agreement was reached. But the discussions at that time helped clarify the advantages and disadvantages of standardization.

We proceeded anyway with some standardization of plant names in our flora. For that purpose we developed a double naming system—employing the scientific name used in the literature and in the specimens' labels, and creating an alphanumerical code with the five initial letters for the family, three digits for the genus (using Dalla Torre and Harms numbers if available), and a numerical sequence for the species and infraspecies. These two parallel components allowed some initial checking for possible mistakes. A preliminary thesaurus was built for the *Flora of Veracruz* on this basis.

Standardization proved useful in dealing with the localities of collection sites. Place names can be quite similar and thus can lead to confusion. We developed a geographical database of Veracruz localities using latitude and longitude as primary keys. This database has been useful for updating old label information and locating with more precision the localities of collection sites on maps. It provides a consistent reference for input of locality data. It also allows us to locate specimens within maps of soil types, climate types, etc. The specimen data can thus be used with geographical information system programs for drawing maps and doing geographical analysis (Olivieri and Granados-Trejo 1990).

We did not attempt any standardization on the remaining 32 data fields in the database: vegetation type, soil, abundance, etc. Exactly what was written in the labels (whether in English or Spanish) filled those fields. We believed at that time, and even more today, that any interpretation of the information on herbarium label information was subject to errors and/or misunderstanding. This was an important decision at the initial stages of the project. Today it is possible to go back to the first captured labels and find the original information. It is the user who will decide what sort of interpretation of this label information will be appropriate for his or her project.

This decision did not mean that we were uninterested in standardization. On the contrary, we have watched these efforts with great interest and will gladly adopt any international standards that are agreed upon. We will use them to develop conversion files for possible users. But we will keep the original label information as a reference point. We have urged collectors to be precise and consistent with others in their descriptions on labels.

*(b) Computer-assisted identification* Computer-aided keys have also been developed using the programs of Richard Pankhurst (Pankhurst 1986; Allkin 1981). Because these keys are not restricted by the nature of print, traits of a specimen can be evaluated in the order desired by the user rather than the order determined by the key's author. The advantages of this approach has been well understood by the many users of the programs in the main headquarters of the *Flora of Veracruz* in Xalapa.

*(c) Quick, easy and wider access* The almost 20 years of routine use of the *Flora of Veracruz* databases for curatorial, research, and informational purposes demonstrate the quickness and ease of informational access. In the past, since few people owned computer hardware, the availability of the information in the computer databases was restricted to a few institutional users. But with the advent of microcomputers and user-friendly software, the information can easily be made available to a

much wider audience. The *Flora of Veracruz* database occupies some 50 megabytes of disk space (including the data themselves and the indexes for quick access to individual records). Previously the database could only be housed on a mainframe or a mini. Today the full database can be placed on a microcomputer with a large enough disk drive. Furthermore, parts can be sent to individual researchers for use on their personal computers with their database programs. We feel that today, more than ever, the use of this approach is the only hope for the future.

## The need to reach a wider audience

But reaching just the professional botanists with this information is not enough to meet the challenge of a disappearing biodiversity. Improving the productivity of the professional botanist is an important goal. But enhancing the professionalism of the amateur and increasing the understanding of the general public about the importance of preserving biodiversity is also vital. The current format of standard floras are not user-friendly (witness the popularity of popular floras). But the totality of the information is important. The question is how to provide a format that satisfies both trained professionals and other users.

We know there is a great need to stimulate research on the plant resources of the tropics. We also must disseminate what we know about tropical plant resources and the main challenges and opportunities they offer. We have to provide information to users and we have to stimulate new users of this information.

### *1. The real status of knowledge*

Here is a concrete illustration of this issue. Suppose a non-botanist from a little town in southern Yucatán Peninsula wants to identify some local plants of economic potential and find out what is known about them. Most likely he will not have the foggiest idea what to do. Almost no school or book in Mexico teaches one how to find answers to such questions.

But suppose our user is an exception and knows about herbaria and systematists. His best first move would be to send his collections to a local herbarium and ask the taxonomist there to identify the plants. The problem is that there is no local herbarium. The closest one is in the city of Mérida, but it has almost no collections from that region, no botanical library and only one taxonomist. The result is easy to predict. The user is not served and is disappointed. Probably he will never try again to formally identify plants and will thus be unable to compare local (traditional) and current scientific knowledge.

This imaginary example is, unfortunately, the rule in developing countries. There are very few local taxonomists and herbaria available. If floras exist, most are written in English. It is clear that they were not written for the local users who need them the most. Furthermore, teachers in schools have no educational materials or means by which to inquire about native plants. In rural communities, students may have traditional knowledge of plants, but they have no way to link theirs with that of modern science. People may be bombarded with newspaper and TV reports on the extinction of plants, the danger of massive deforestation, and the importance of biodiversity, but no connection is made to their day-to-day activities.

*2. New herbaria: a solution?*

Clearly we need to change this situation. Unfortunately, there is a widespread belief that the only way to solve the problem is to establish new herbaria. The idea is good but the reality is that many of these new institutions never become big enough to be meaningful. They are usually underfunded or not funded at all and their resources are minimal. So, the result is poorly curated collections that provide very limited service to users.

These new local herbaria have not countered an unfortunate, but strong, trend in botany. Important floristic research has become more and more concentrated in fewer and fewer institutions of the developed world. Even though some of these institutions have obtained a high level of excellence, they do not reach local users in the countries that need the information the most. Efforts have been made to alleviate this problem, but they are negligible.

The production of floras is strongly linked to this problem. Few floras of developing countries are actually developed and published in the developing countries themselves. Instead, more and more floras of developing countries are now under way in the larger herbaria of the world, i.e. those in the developed world.

The great paradox of this is that the worldwide need for stronger botanical institutions to provide the necessary information for biodiversity conservation accompanies an alarming trend of extinction of herbaria and systematic programmes everywhere. Something very wrong is happening.

There are tropical countries without a major herbarium. There are countries or immense regions without even a plant taxonomist. The future of these countries is based mainly on their plant resources. Yet they don't know what they have and they do not have the basic facilities to find out. International pressure exists for the conservation of biodiversity, but no programmes exist to help developing countries identify and learn how to identify their resources.

## The basic problem: information access

A major reason for these problems, in addition to lack of funds, is the lack of more efficient means of access to the available information. If we improve that access, we would also improve the basis for the identification of new species as well as known ones. We would also improve the process of learning more about a species' distribution, abundance, and threats to its existence. But the key is the facilitation of the process of identification of unknown plants.

### *1. The promise of new technology*

During the development of the *Flora of Veracruz* project, we struggled with these problems: the establishment of a new herbarium, the development of databases and geographical information systems. We really wanted to include pictures. But at that time, there was no easy and financially reasonable way. But we knew from experience that something different must be done. Pictures provide an immediacy of description not available from textual depictions unless the reader has experience.

The introduction of technology allowing the storage of large number of images allows a new approach—the combination of floristic databases with videodiscs. In the past two years we have experimented at the University of California in the Plant Resources Information Laboratory with the possibilities of using the textual databases with optical storage media (Gómez-Pompa and Plummer 1990). We have developed a demonstration videodisc for the *Flora de Veracruz*. Included among the images on this disc are: herbarium specimens, colour and black and white photographs, slides, and drawings (Fig. 9.1). We are also developing a video database of orchids of Mexico using optical disc technology, in collaboration with the Orchid Society of Mexico. The main objective of that project is to create a special programme for the orchids of Mexico. We have also been developing several other video projects: the Zamiaceae of Veracruz in collaboration with Andrew Vovides of the Institute of Ecology of Mexico, the trees of the Maya area in collaboration with several institutions working in that region, as well as other taxon-oriented projects. For many of these projects, we are producing interactive programs for education and research, an approach we call Q'Taxa (Gómez-Pompa and Campos 1989). This video was shown at the Delphi symposium.

Our experience with these projects has convinced us that optical storage technology will be an extremely useful and economical solution to the problems of plant identification and plant information. Most importantly, it is a solution in which developing countries can actively participate in both its production and its usage.

Floristic databases that include laser videodisc technology provide a

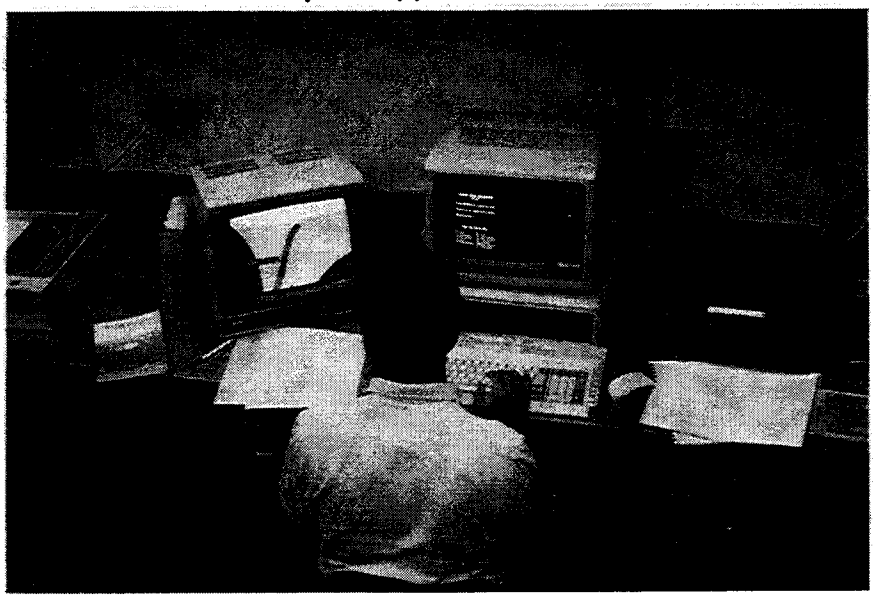


**Fig. 9.1.** A video flora workstation with computer, laserdisc player, and separate monitor.

partial solution to the limited accessibility to information on plants for identification purposes and to information in general on plant resources. Laser videodiscs have two properties important for a flora: huge image storage capacity and interactivity. One side of a twelve-inch videodisc can hold 54000 different images (drawings, photographs, microscopy, etc.), that is, far more than a book. Each image on the disk can be brought to the screen in less than a second and at will, that is, interactively (by commands to the videodisc player from a computer program or remote unit). This exciting technology may provide temporary or permanent alternatives to the publications of conventional floras (Fig. 9.1).

Herbarium specimens provide the best register of the existence of a plant in a region. They are and will be the central tool in systematic research. Unfortunately they are limited in distribution and availability, especially those collected many years ago that have been designated as the types of the nomenclatural species that we use. These specimens are crucial in the identification process. For that reason they become the principal obstacle in the identification of plants. Access to them is limited, despite the good efforts of several organizations that make photographs available to users. Very few herbaria in developing countries have good phototype collections as they are expensive, even in microfiche.

While a well-curated herbarium is the best tool for plant identification,

we see in the future a very different herbarium: the electronic herbarium. Any institution that may decide to produce a video flora will be taking the first step into a new network of floristic projects around the world and into the establishment of the electronic herbarium.

This new facility will occupy less space and will have in its library video floras from many parts of the world and also bibliographic information on the plants. The new exchange will include videodiscs of the local projects in addition to specimens. The electronic herbarium will require at minimum a videodisc player attached to a computer (whether a IBM-compatible PC, MacIntosh, Sun workstation, or whatever).

A video flora can be a very powerful tool in the identification of plants. The great advantage is that the identification can be done almost *in situ*. Several hundred thousand images of specimens and species along with accompanying data can be accommodated in a few laserdiscs. The images and data one needs for identification or other information purposes can be at one's finger tips in a school classroom, or library, even at a field station – anywhere where electricity is available (through existing lines, gasoline generators, or solar panels). The same disc can be used for many different purposes including education, ecological research, etc.

The intent of a video flora is to provide an understanding of the flora in a certain region for the expert well versed in the plants there, the botanist not so familiar with the plants, and the interested non-botanist.

## The challenges

It is obvious that the value of a video flora disc will be related to the precision and accuracy of its information. Images based on actual specimens and slides of living plants from which those specimens were collected have a higher value than those that do not have such a voucher. And here is the Achilles' heel of this approach. The gathering of data and images has to be done in co-ordination with an active field research programme. Also, the recording of information about each image (whether used in the final videodisc or not) into a computer database is essential for quick and efficient production of the video flora.

This effort may take a lot of time and energy. This work should be done by local scientists who will produce their own local image banks that may be offered in exchange for other similar projects, bibliographic information, specimens, etc. Through emphasis on local effort and exchange, we are encouraging organized field work for a precise objective: the production of a video flora. Videodisc players will in a few years be as common and as popular as videotape players. Computers are everywhere. The local centres will be enriched with information and local collections. Larger institutions will have duplicates of the laserdiscs and the information. And

most importantly this new accessibility will promote more activities in the study of local floras by all kinds of people from school children to amateur botanists.

One challenge is to produce programs at a reasonable price. A lot of money can be spend on video production. The top-of-the-line equipment is expensive. Commercial video production firms charge very high fees for studio and personnel time. But a useful, good-quality videodisc need not compete with television commercials in cost. There are avenues for sharing production resources and expertise that can keep costs reasonable.

Another challenge lies in providing a user interface that allows the user to explore the images on his or her own. People need quick, easy and self-initiated comparisons between images for understanding. The user should be able to pick her or his own list of images for quick and easy comparison. But again the interface need not be an individual effort.

We see the video flora as an ideal solution for the crisis in plant exploration and research for the following reasons:

1. The cost and resources for producing satisfactory video floras are within the capacity of most countries of the world and the agencies and foundations funding natural conservation and resource management.
2. The cost of providing the equipment (videodisc players, CD-ROM drives and even IBM-compatible or MacIntosh computers, if necessary) for the research institutes is also within the funding capacity of world conservation agencies and foundations, not to mention, governments—around US $7500 for remote places without any such equipment.

The challenge is in the acceptance of this approach by the botanical community and by the granting agencies: to encourage, support and fund video flora projects.

At University of California at Riverside, we are developing a project for the Maya area that will include visual and textual floristic databases, GIS, and identification programs that will be used by a network of small institutions in the Maya area. We are also proposing to experiment with computer-mediated bilingual courses for the region that will include faculty and students from California and Mexico connected through BITNET to learn about the Maya area, its natural resources, and its people. Our video flora approach will be an important component of this project. There is a great future for these approaches on a global scale. The hardware and software is already in place and at a reasonable cost. What we need is ideas for their use and funds to develop the finished projects.

## A project to be developed: a world's flora of protected areas

It seems to us that there is a great need to develop as soon as possible a checklist of the plants protected in the network of protected areas of the world. This project should be done through an international agreement between governments, botanical institutions, and field scientists concerned with protected areas. A minimal database has been suggested for this project (Gómez-Pompa *et al.* 1988). It is our feeling that a project of this magnitude could be the best initiative to support a global plant information system. The use of the video-flora approach for such project could be a extremely important decision. It seems to us that the time is right for an ambitious programme of this kind. The world needs to know what are the biological assets protected in order to plan future reserves. No one could do it except the systematic community.

## A final word

We need to change. We need to agree upon some standard formats. But the most important thing that has to change is the acceptance to share databases and programs as if they were publications. The more people read them and use them the better. We should not confuse the business world of software and hardware with our own business: to learn and to promote learning about nature.

## References

Allkin, R. (1981). *Métodos para la identificación de las familias de Angiospermas de Veracruz*, Informe interno de INIREB, No. 8130023. Instituto Nacional de Investigaciones sobre Recursos Bióticos, Xalapa, Mexico.

Gómez-Pompa, A. and Campos, E. (1989). *Q'Taxa*. A video tape produced at University of California at Riverside. Copies available on request (include cost of video tape and mailing; state format required).

Gómez-Pompa, A. and Nevling, L. I. Jr. (1988). Some reflections on floristic databases. *Taxon*, **37**, 764–75.

Gómez-Pompa, A. and Plummer, O.E. (1990). Video floras: a new tool for systematic botany. *Taxon*, **39**, 576–85.

Gómez-Pompa, A. and Squires, D.F. (1969). A symposium on information problems in the biological sciences. In *Abstracts of the Symposium on Information Problems in the Biological Sciences*, pp. 7–9. Publicaciones Especiales del Instituto de Biología de la UNAM, No. 1, Mexico City.

Gómez-Pompa, A., Castillo, S., Plummer, O. E., and Soto, M. (1988). *Biosphere reserves: a minimal database*, a report to the International Union for the Conservation of Nature and the World Wildlife Fund. IUCN, Gland Switzerland.

Olivieri, S. and Granados-Trejo, M. (1990). *Sistema de Información Geográfica*, programs and manuals. INIREB/CI/UC Riverside. (Distributed by Conservational International, Washington, DC.)

Pankhurst, R.J. (1986). A package of computer programs for handling taxonomic databases. *Computer Applications in the Biosciences*, **2**, 33–9.

Wirawan, N. (1989). The Flora Malesiana symposium and workshop, Leiden. *Wallaceana*, **57**, (September), 8–9.