

WHAT IS BIBLIORCHIDEA?

RUDOLF JENNY

Moosweg 9, 3112 Allmendingen, Switzerland

RJe@io3s.com

Introduction

Literature is still playing a major role in science and research, on one side because it is a documentation of research already done and a demonstration of the results. On the other hand it is also a record about recent research. At least for the time being, literature search is a part of any scientific work or project. Whoever has done literature search knows about the difficulties to reach an overview of the literature in connection with a project in an acceptable time and with acceptable effort. Computer technology today provides us with a large selection of very helpful tools to limit time and effort for literature search.

As explained above, literature still plays a very important role in science. In this context the term literature should not be defined too narrow, we have to accept a very broad collection of publications as literature. In botany – and in orchideology as a division of botany – we find:

- Scientific periodicals, occasionally with articles about orchids
- Scientific periodicals in the botanical field, occasionally with articles about orchids
- Orchid periodicals, with at least partially scientific content
- Society publications (Orchid Societies)
- Dissertations
- General floras or orchid floras
- Proceedings, abstracts and reports of congresses and symposiums
- Catalogues of all kind
- Travel and expedition reports
- Textbooks and basic research publications
- Bibliographies and Biographies
- Correspondence and letters

Botany as an independent scientific discipline is not very old. For centuries, botany was together with other natural sciences like zoology or geology always

a part of medicine or part of a general study in natural sciences. The oldest European botanical literature was always very closely connected with medicine and pharmacology. In this time orchids were treated as a mere curiosity of nature and, in a limited number as part of pharmacological publications like herbals. Only a few publications devoted to orchids alone are known before 1800, after this time we see a fast increase in the number of publications devoted only to orchids. Today we have more publications about orchids than about any other plant family. Because of the more and more interdisciplinary connection of botany with other natural sciences like chemistry, biology or zoology the content of publications about or in connection with orchids has also become much broader. Like in many other sciences also in botany the pure generalist does not exist anymore. One result of the ongoing specialization in botany is the fact that the term botanist is no longer a synonym for classification, systematics or taxonomy. With this the existing literature becomes less and less manageable, and it is almost impossible to avoid repetitions in publication and research. For a good part, this is the result of the fact that today it is almost impossible to keep track of the enormous numbers of publications, in spite of the fact that we have tools like the internet at hand. One should think that based on the available data processing technology it should be possible to solve this problem in a simple way by building up databases or computerized bibliographies. This is not the case and the reason can be found in the philosophy and structure of such a database. We have to consider first the enormous variability concerning the content of orchid literature:

- Classification, systematics and taxonomy
- Nomenclature
- Genetics, molecular biology, DNA-analysis, enzymatics
- Anatomy and morphology
- Phytogeography, distribution, mapping

- Evolution, speciation, population dynamics
- Interactions plant - animal, pollination ecology
- Symbiosis, mycorrhiza
- Phytochemistry, fragrance, pigmentation, pharmacology
- Vegetation biology, habitat, ecology
- Physiology
- History
- Hybridization, breeding, commercial aspects
- Culture
- Propagation, micropropagation, tissue-culture
- Protection and conservation

Naturally this list is not complete by any means, but it shows the very complex interdisciplinary connections in orchideology and the enormous spectrum of contents of publications about orchids.

An electronic library or literature database may include a very wide selection of publications from all the above mentioned fields, not only about orchids but about botany in general. To use a metaphor: such a database would be very broad but not deep, it would be a rather flat disc with a large diameter. The lowest level in such a database would be the family *Orchidaceae*. Such a database would produce a large number of general information, but it would be very inefficient for an orchid specialist because a search for a combination of keywords like „molecular biology“ **or** *Orchidaceae* would end up in a vast and confusing number of citations. The search for the combination „molecular biology“ **and** *Orchidaceae* on the other side would end up in a small but incomplete number of citations. The contrary of such a „general“-database is a „special“-database, in this case also it would contain publications from all the above mentioned fields but only such documents in connection with *Orchidaceae*. To use the same metaphor: this database would be like a cylinder, very deep but with a limited diameter. Such a database is build especially for users interested in *Orchidaceae*, for other users the result of a search would be to specific.

The structure of a computer-based bibliography is variable in detail but the overall principle is always the same: the possibility to search for literature based on different criterions. The question whether the entire publication, the summary of a particular publication or simply the citation together with keywords are avail-

able in a database, depends only on the availability of the literature itself and the situation concerning copyrights. Older literature is already in the public domain, the question is whether the effort to scan such publications is proportional to number of accesses by the users of the database. To scan a very old book will be difficult by any means and the effort is high, it would not make sense to spend time and money for a very few interested users only. The size of a computer-based database is basically a function of two parameters: first the human resources – time and financial support – of the institution which is maintaining the database, and second the availability of publications fitting in the frame-work of the database. Especially the financial point is the limiting factor for size and completeness of an electronic database. The more specialized this database should be and therefore the less potential users one may expect, the more difficulties an institution will have to obtain the financial resources. This criterion is - at least for the time being - independent from the available technology. The question about the completeness of a computer-based literature database in science is therefore easy to answer: the narrower the definition of the content the higher the degree of completeness, the broader the definition, the less complete the database will be.

Another criterion to judge the quality of a database is the strict neutrality concerning the importance and quality of the included publications or documents. It is a common place that those who are maintaining a database will have almost certainly their own ideas about importance and quality of the content of publications they include. This idea is also almost certain different from the opinion of the users. It is therefore paramount to avoid any valuation of publications, if such documents fit in the definition of the database, they have to be included. It is strictly up to the user to make an own selection and judgment of the content of a given document. A third criterion is the consistency of a given database, it is important that the internal organization of a database is consistent and that also the procedure of adding new documents guarantees consistency. In other words, a given document should have the same keywords and the same form, independent when the entry was included in the database and by whom.

In order to understand what system and philosophy

a database should follow, we need to understand the needs and wishes of the user. Everybody who was sometime or other in the position to do literature research for an own project or publication, knows exactly how time consuming and often expensive this task can be. Such a process can be divided in two very clearly separated phases. The first step is the providing of a list of publications or citations in connection with the target of the planned work, the second step is to obtain the important publications in full insofar as they are really needed.

The first phase often is extremely time-consuming because such information has to be compiled from different physical sources or electronic databases. With exception of BIBLIORCHIDEA there is no other comprehensive database or bibliography of literature exclusively about or in connection with orchids in all aspects existing. An example: the Index Kewensis covers all taxonomic first descriptions of plants including the members of the orchid family, but there is no information about available monographs or revisions of a given orchid genus. Dissertations are covered by specialized databases of their own, and information about publications from interdisciplinary areas like molecular biology are not or only in parts included in databases about botany. The search for literature will therefore end up in a more or less long excursion through different computer based databases or printed sources. Unfortunately most of those databases will have a different system, a different user interface and different search engines.

The second phase, the acquisition of the „real“ literature based on several lists of citations, is also often rather difficult and time consuming. Usually public libraries or university libraries are delivering on request exactly what the customer is ordering, nothing more and nothing less. If the citation is wrong or incomplete the customer will get wrong or incomplete response, occasionally the library will ask for more detailed information. The mere number of definitely wrong literature citations in publications is amazing and frightening; the range spreads from invalid or incomplete abbreviations to wrong volumes, wrong authors and wrong page numbers. Obviously the process of search for literature has become so time consuming and expensive that in order to save time

authors are copying citations from other sources without ever have seen the literature itself. It is an open question whether this is a scientifically acceptable way to work, but the example shows the problem for a scientist to collect the necessary literature for a given project in acceptable time. Even if a good literature collection is at hand, the problem is not solved, the search for certain things and without clear citations in available literature is time consuming too. It seems to be clear that today no library can employ an orchidist in order to handle orders for orchid literature, and there are many other plant families with exactly the same problem. It is also clear that library staff cannot spend time to check in detail all unclear or incomplete orders from customers, there is but limited time available, if it cannot be spent the order goes back to the customer with a respective remark. There is one consequence out of all these facts: search for literature in an acceptable time and with acceptable effort can be done only by using a specialized database with a library in the background in which we find physically all the documents or publications included in the database and with a staff who is specialized in this area. The combination of library, database and specialized staff is paramount.

For the people maintaining the database it is important to know what the potential user needs. Also in this case the spectrum is very broad, from a simple search for the correct spelling of a certain epithet to a search for literature as basis for a monograph or dissertation almost everything is possible. The structure of the database should ensure that questions from all over this area can be answered. Because no database is complete, it cannot be expected to get a complete list of publications about a certain issue, but the list has to be complete enough for a start.

BIBLIORCHIDEA

Based on the fact, that the time available for literature search is limited, the project BIBLIORCHIDEA has been developed over a period of about 18 years. The inner structure of this computer based bibliography and the story of its development is a good example to show how a „special“ database is build up. At the beginning it was a very simple structured list of available books and periodicals in an already rather large private orchid library, the main target was to ensure that the same book was not purchased or ordered twice.

There was no connection of the entries by selected keywords, the target was only to get fast information whether a certain publication was available or not.

In a second step, this list was integrated in professional and special software for library management. At this time the used software LIDOS was one of the most sophisticated and powerful tools available in this area. Very soon it became clear that the capacity of the used software would allow a much better and finer definition of the integrated entries or documents, and with this a much deeper and further use would be possible. It would result in much more than just a list of publications available from a library. So all the necessary information to order or to find a given publication has been added as well as information about author, co-author, year of publication, title, editor and publisher. Another point was the selection of the used standards on which citations would be based. For plant names *Index Kewensis* is the standard, for single publications (books) it is STAFLEU's *Taxonomic Literature*, for periodicals and journals it is BPH and BPH-supplement and for everything in connection with herbariums it is *Index Herbariorum*.

In a next step a keyword catalogue was build up, this collection of keywords would allow a search for literature after its content. Right from the beginning a hierarchically organized structure was chosen, which would make it possible to select the appropriate degree of selectivity for each search. The difference between a search for all publications about orchids in Europe, or orchids in Switzerland or orchids in the area of Zuerich is quite obvious. For Europe as keyword we would end up with virtually thousands of documents, for Switzerland still with hundreds and for Zuerich probably only with some twenty documents. This hierarchical structure allows the user a very specific search with a manageable and clear number of answers or documents. Today there are six different levels in the keyword catalogue, an example will show this:

1 th level (main keyword)	Geography
2 th level	North America
3 th level	USA
4 th level	Florida
5 th level	Everglades

Very soon it became clear that with this structure it would not be possible to integrate periodicals or jour-

nals. Hence in a next step all entries concerning periodicals have been removed and replaced by the article in the periodical itself. In order to integrate the correct citation of a particular article a new submenu or field in the entry menu had to be created and the title of the periodical was consequently integrated in the keyword catalogue. With this a search for a particular article and the search for all articles in a particular journal has become possible. This step of integrating articles was connected with enormous effort. Between 1841 and 1902 the well-known journal *Gardeners' Chronicle* alone contained not less than some 12'000 articles about or in connection with orchids, the reports about the sessions of the Orchid Committee of the Royal Horticultural Society not included. Today *BBIBLIORCHIDEA* contains about 120'000 articles from about 1'400 different journals and periodicals. Some of these journals are integrated completely, that means from volume one up to the recent number or volume with all articles (e.g. *Orquidea (Mex)* and *Orquidologia*), of others all articles in connection with orchids are integrated (e.g. *Selbyana*, *Botanical Leaflets Harvard University*) and of some only the known articles about orchids are integrated (e.g. *American Journal of Botany*). The main problem here was access to the primary literature, some of these journals are rather difficult to obtain because they are old or because they are not very widely circulated in libraries. Together with the titles of the journals the keyword catalogue increased to a number of some 25'000 keywords.

In the same time another submenu or field in the entry menu was introduced, in the field „species and below“ we find an alphabetical list of all new taxa below generic level described in the particular publication or document. This part is in fact something similar to the *Index Kewensis* but the information is neutralized. That means the information is not whether a certain new taxa is valid or not according to the rules of botanical nomenclature or whether it is a synonym of something else, the information just states that the particular author had described this particular taxa in this particular publication. Again, it is up to the user to evaluate the information. Right from the beginning also varieties, formae and subspecies have been included; these names are not included in the older volumes of the *Index Kewensis*. In the meantime

about 65'000 taxa below generic level are included in Billbiorchidea, with this an estimated rate of about 80 % of all described taxa below generic level in the orchid family is already accessible. Naturally also all taxonomic first descriptions of entities between species level and subfamily (e.g. genus, subtribe, tribe, hybrid genus) are included in the database, these entities are integrated in the field „above species“.

The fast growing keyword catalogue showed that in some cases an explanation of the keyword was necessary. Especially in some categories of keywords like titles of periodicals or individual personal names, an explanation of the keyword becomes important. The reason for this lie in the fact that by definition keywords have to be short and clear - no sentences - and that, wherever abbreviations are used, the complete, not an abbreviated form has to be available also. The example of the BACKHOUSE dynasty in England shows very clear why such explanations are important. The names of all three members of the BACKHOUSE dynasty are included as keywords, all three have the name James BACKHOUSE, the only difference is the date of birth and death. To avoid mistakes, the commentary to those three keywords explain exactly which James BACKHOUSE was father, son or grandson. These commentaries are accessible through the keyword catalogue. Especially important are the commentaries in the categories hybrid genera (parents and valid RHS-abbreviations), individuals (personal data), book series and periodicals (abbreviations after BPH and information about changed titles). The commentary of the title of a given periodical contains the full title of the journal, the official abbreviation after BPH, information about the time and extent, and information about a possible succeeding and preceding title of the journal, again with full changed title, BPH-abbreviation and extent. With this information a journal is defined in a very clear way, which is important considering the fact that rather often titles of journals are quite similar (e.g. Orchids (AOS), Orchids (South Africa) and Orchids (Australia)).

Another change was the decision to integrate iconographies like the *Lindenia* or *Icones Plantarum Tropicarum* not as a complete and single document, but by plates, that means every single plate was treated as a document of its own. With this decision again the number of entries or documents was increasing dramatical-

ly, the second edition of the field guide of the Orchids of Venezuela, published in 2000, alone added some 1'100 new documents to the bibliography. Together with this increase also the content of information increased, it was now possible to search at the level of single species and to find very fast an illustration of a particular plant. The plate by plate introduction of complex publications like the above mentioned Orchids of Venezuela and *Icones Platarum Tropicarum*, or like *Flora Brasilica*, *Flora Brasiliensis*, *Venezuelan Orchids Illustrated* and many others, was completed after about one year. The result of this task is the possibility to gain much more detailed literature citations.

The very fast development in computer technology and also the availability of better and more sophisticated software in connection with the fast growing importance of the internet were responsible for another decision about the future of BIBLIORCHIDEA. The existing database in its original DOS-based interface was available for interested users for several years under the name LITBUL. In order to keep BIBLIORCHIDEA up to date concerning the large number of new publications and because the structure especially of the keyword catalogue was changed and enlarged from time to time, a simple upgrade for the user was not possible. The only solution was a complete renewal of BIBLIORCHIDEA at least once a year. With this interval the database was in the worst case about one year behind. The process of creating renewals was expensive and not very efficient. Since the value of such a database is measured also on its being up-to-date, it was very important to find a way to maintain it in “real-time” via internet. Beside this, the used DOS-based software was old, it was not possible to get print-outs of a search result in an easy way without data-transfers into word processing programs and it was not possible to use the mouse. This overall unsatisfying situation could be changed only by a fundamental change of the software environment. Consequently the decision was taken to extract all the data from the old software and to put them in a totally new software environment and, consequently make the new form accessible through the internet. Today new entries, corrections or changes in the data are done directly via internet, as a result of this on-time maintenance, BIBLIORCHIDEA is up-to-date all the time.

Naturally also BIBLIORCHIDEA is not complete

and there is much doubt whether it ever will reach completeness, but it will grow continuously. Some orchid journals like *Orchid Review* or *American Orchid Society Bulletin* are not yet completely integrated and from other periodicals we still are looking for missing volumes. On the other hand, many periodicals are already in the library but not yet included in BIBLIORCHIDEA. The limiting factor is not only the time, it is also the availability of the literature; in order to add new entries or documents, we need a physical copy of the original publication, without this we would copy old mistakes and we could not define the keywords. For the next five years we will have about 10'000 new documents per year, of these about 2'500 will be new publications and about 7'500 old publications which have become available in the meantime. With about 150'000 entries we will reach a platform, the increase per year will then be reduced to about 2'500 new publications and about 500 old ones. Probably we will then have the time to reprocess some documents with the goal of a further and finer classification. Especially some of the fundamental works about orchids, like SCHLECHTER's publications in *Feddes Repertorium*, we would like to divide in smaller parts. Today BIBLIORCHIDEA contains about 140'000 documents, included in this number are articles from all kind of periodicals, books or „single publications“, catalogues, dissertations, checklists, manuscripts and iconographies. All these documents, as far as they have been published, are included in a form with enough information to order them through a public or scientific library. All of them are also represented in our library as physical copies. One of the hopes for the future is that authors all around the world would realize that the best way to make their own publications known would be to send us a copy in order to add it to BIBLIORCHIDEA as fast as possible. This is especially

important for publications which are not widely distributed, like dissertations.

The actual form of BIBLIORCHIDEA as it is accessible through internet (www.Bibliorchidea.org), will allow the user different kinds of search or also the connection of different search methods. These are

- Direct search for author and co-author
- Direct search for the year of publication (selection direct or in a time-window)
- Search with text-input in the fields Title, Literature quotation, Editor, Publisher, Above species and Species and below (free-text search)
- Direct search for new descriptions in the respective fields Above species and Species and below
- Search for keywords by direct selection from the keyword catalogue as single keyword or in connection with other keywords by using the connecting terms and / or (Boolean connections)
- Enlarge the result by using one of the above mentioned methods.
- Restrict (decrease) the result by using one of the above mentioned methods (decrease to and Decrease by – functions)

Besides the search mechanisms, the software naturally allows the sorting of results by different criterions and the printout as list of documents or as single document with all the detailed information. The important addition is the fact that all documents the user can find in BIBLIORCHIDEA are also available as physical copy, hence a very fast access is guaranteed.

According to the very fast technological development especially in the information technology, it is extremely difficult to guess what development a database like BIBLIORCHIDEA will see in the next years. Certainly BIBLIORCHIDEA will remain a most important tool for everybody who need orchid literature for profession or hobby.

Rudolf Jenny is chemist by training and director of a company active in the industrial application of ozone. He is interested in the taxonomy, pollination ecology and history of orchids, especially in the subtribes *Catasetinae*, *Huntleyinae* and *Stanhopeinae*. He is author of treatments of several genera in those subtribes, especially *Stanhopea* and *Gongora*. He owns one of the most complete orchid libraries in private hands and maintains the literature database BIBLIORCHIDEA. Beside this he is regularly publishing about orchids in different periodicals.