

Zeitschrift:	Candollea : journal international de botanique systématique = international journal of systematic botany
Herausgeber:	Conservatoire et Jardin botaniques de la Ville de Genève
Band:	45 (1990)
Heft:	2
Artikel:	Concept and use of an integrated database system for the Chaco : application to a preliminary checklist : contribution to the study of the flora and vegetation of the Chaco : II
Autor:	Zellweger, Catherine / Palese, Raoul / Perret, Patrick
DOI:	https://doi.org/10.5169/seals-879720

Nutzungsbedingungen

Die ETH-Bibliothek ist die Anbieterin der digitalisierten Zeitschriften auf E-Periodica. Sie besitzt keine Urheberrechte an den Zeitschriften und ist nicht verantwortlich für deren Inhalte. Die Rechte liegen in der Regel bei den Herausgebern beziehungsweise den externen Rechteinhabern. Das Veröffentlichen von Bildern in Print- und Online-Publikationen sowie auf Social Media-Kanälen oder Webseiten ist nur mit vorheriger Genehmigung der Rechteinhaber erlaubt. [Mehr erfahren](#)

Conditions d'utilisation

L'ETH Library est le fournisseur des revues numérisées. Elle ne détient aucun droit d'auteur sur les revues et n'est pas responsable de leur contenu. En règle générale, les droits sont détenus par les éditeurs ou les détenteurs de droits externes. La reproduction d'images dans des publications imprimées ou en ligne ainsi que sur des canaux de médias sociaux ou des sites web n'est autorisée qu'avec l'accord préalable des détenteurs des droits. [En savoir plus](#)

Terms of use

The ETH Library is the provider of the digitised journals. It does not own any copyrights to the journals and is not responsible for their content. The rights usually lie with the publishers or the external rights holders. Publishing images in print and online publications, as well as on social media channels or websites, is only permitted with the prior consent of the rights holders. [Find out more](#)

Download PDF: 12.01.2026

ETH-Bibliothek Zürich, E-Periodica, <https://www.e-periodica.ch>

Concept and use of an integrated database system for the Chaco. Application to a preliminary checklist Contribution to the study of the flora and vegetation of the Chaco. II.

CATHERINE ZELLWEGER, RAOUL PALESE
PATRICK PERRET, LORENZO RAMELLA
&
RODOLPHE SPICHIGER

ABSTRACT

ZELLWEGER, C., R. PALESE, P. PERRET, L. RAMELLA & R. SPICHIGER (1990). Concept and use of an integrated database system for the Chaco. Application to a preliminary checklist. Contribution to the study of the flora and vegetation of the Chaco. II. *Candollea* 45: 681-690. In English, English and Spanish abstracts.

The Conservatoire et Jardin botaniques de la Ville de Genève (CJB) are proceeding with their study on the flora and vegetation of the Chaco. For this purpose, a preliminary checklist based on the Geneva Chaco collections is undertaken. This inventory is established using the integrated database system of the CJB. The concept and structure of this database are presented as well as the available computer equipment. The elements of the integrated database which are useful for the checklist are analysed. The distribution map related to the checklist is also given.

RESUMEN

ZELLWEGER, C., R. PALESE, P. PERRET, L. RAMELLA & R. SPICHIGER (1990). Concepto y utilización de una base de datos integrados sobre el Chaco. Aplicación para un catálogo preliminar. Contribución al estudio de la flora y de la vegetación del Chaco. II. *Candollea* 45: 681-690. En inglés, resúmenes en inglés y español.

El Conservatoire et Jardin botaniques de la Ville de Genève (CJB) prosigue el proyecto sobre el estudio de la flora y de la vegetación del Chaco. En este contexto, se comienza un catálogo de táxones tomando como base las colecciones chaqueñas depositadas en Ginebra. Este inventario se realiza después de la elaboración en CJB de una base de datos integrados. Se presenta el concepto y la estructura de esta base de datos así como el material informático de que se dispone, al mismo tiempo que se analizan los elementos de la base de datos integrados útiles para este catálogo. Se presenta también un mapa de la región considerada.

Introduction

In 1987, the Conservatoire et Jardin botaniques (CJB) de la Ville de Genève undertook a research programme concerning the flora and vegetation of the Chaco (RAMELLA & SPICHIGER, 1989). This project was in addition to that of the "Flora del Paraguay" carried out from 1982 onwards in collaboration with the Missouri Botanical Garden of St. Louis. At the same time, the CJB received data processing equipment, as well as human resources, which enabled them to envisage the development of a database system. In this field, the needs of the CJB were varied, ranging from collections management, the development of tools for the scientific projects, with, in particular, the creation of specific databases for each of them, not to mention office automation.

The considerable thought given to the matter resulted in a concept which would enable integration of the data and their sharing out into as many applications as necessary for the projects. Among the latter, the study of the flora and vegetation of Paraguay and the Chaco derives from the fact that Geneva has historical herbaria relating to this region (SPICHIGER & MASCHERPA, 1983). This situation, to which can be added the important resources of the CJB library, led to the setting up of an application of the integrated database system for a preliminary checklist of the flora of the Chaco. This checklist is designed as a working tool which will evolve both in structure and contents. But right from the outset it was agreed that the information included would be rigorously checked and standardized. The aim in mind was to ensure that, in the medium term, it would be possible to have access to a true information system on the flora of the Chaco.

Concept and use of the computer tool

The first use of data processing at the CJB goes back to 1979 (MASCHERPA & BOCQUET, 1984). The task of the computer, at that time, was related to the collections management and local processing of data concerning a specific project (BOCQUET, 1982). Computerization at the CJB followed a classical development, and today offers the privilege of communication and integration within the institution. This transformation, which was initiated in 1986, leads to major changes in respect of the working methods; it is therefore of interest to describe this new computer tool.

Hardware and software

The CJB are equipped with a Local Area Network (LAN), of the Ethernet type, to which the computers and peripherals are connected. Two computers of the µVax range of Digital Equipment Corporation (DEC), one devoted to telecommunications and to office automation, the other to the databases, carry the major part of our applications. Two workstations, of the Vaxstation 3100 type, configured in cluster, are used for development purposes. Two IBM-PC compatible personal computers (PC) are used for automatic mapping as graphic terminals and as communications ports on to the MS-DOS world. The network at present comprises 40 terminals and several shared printers. An essential element of the LAN is the MCS Powerview 10 (Compugraphic) photocomposition station. By means of a conversion procedure, this machine directly processes the documents compiled by text processing or the files obtained by querying the database.

The µVax and workstations function with the VMS (DEC) operating system. The All-IN-1 (DEC) integrated software for office automation offers all the necessary functions of communication (electronic mail) and text processing. At this point, it should be mentioned that the LAN of the CJB is connected to the network of the University of Geneva which offers access to the international telecommunications networks.

The Database Management System (DBMS) used is Rdb/VMS (DEC). It is built on the relational model and provided with an SQL interface. All of the applications necessary for access to the database were developed at the CJB, using Pascal and TDMS (DEC).

Concept of integrated database system

The databases used in taxonomy have been classified and described according to their contents (BISBY, 1984). We shall apply this classification to the classes of data relating to taxonomy and define them as follows:

- **Geography:** gathering localities, data relating to the distribution of the taxa.
- **Specimens:** all the data relating to gatherings, to samples, collections.
- **Persons:** all the data related to collectors, determinators, authors and names authorities.
- **Nomenclature:** complete names, with authors, place of publication of the protologue, typification, comments.
- **Taxonomy:** accepted name and synonyms, morphological descriptions, habitat, etc.
- **Bibliography:** bibliographical references, thesauri and access keys.

Flora-related projects call for a synthesis of this information. The integration of the data relating to these various classes of data, in a single tool, using a single collection of data, is necessary. Several realizations offer a good level of integration; we can, for example, quote the database of the "Flora of Ecuador" (FROST-OLSEN & HOLM-NIELSEN, 1986) or "Tropicos" (CROSBY & MAGILL, 1989). But the complexity of the data involved, especially in taxonomy, runs into the problem of the classical architecture of commercial software, and calls for expertise in the structuring of the database and programming (ALLKIN, 1988). The development of new technologies will perhaps bring simple solutions, which are user-friendly and less expensive (SKOV, 1989).

The design of the CJB's database was conceived in order to achieve two levels of integration: that of the classes of data and that of the institution's projects.

The first objective is therefore to offer several views for using the unique collection of data according to the scope of interest: **Geography, Specimens, Persons, Nomenclature, Taxonomy, Bibliography**. The second aim is the integration of the data concerning all of the projects of the institution in the same collection of data. The notion of usage context is defined as being the specialized use of the database by one or several persons within the framework of a project, often associated with a geographical context. A usage context delimits a data subset.

The advantages of this design of integrated database are mainly the following:

- sharing of the data in real time
- reduction of redundancy
- creation of shared thesauri
- comparison of expertise
- efficiency of validation and standardization of the data.

We should stress at this point the importance of the characteristics of the equipment and software available to us when the system was being designed. First, the centralized information system used at the CJB favoured the creation of a monolithic database, which meant that we avoided the difficulties linked with distributed databases. The applications were developed on a made-to-measure basis to prevent non-specific access to the database and to enable users to go straight to the usage context they desired. Secondly, the multi-user DBMS, which has a strong management for accesses, for data integrity and transactions enabling synchronization of operations, deals with the problems of competition for data. Its relational architecture ensures development in a progressive manner, with great ease of database restructuration.

The dual integration of the classes of data and the projects can be functional and durable only if the database is administered. Administration is particularly critical in respect of:

- validation of the data:

All information which is added to the database is provided with a record containing the name of the creator and the date of creation. This information is "doubtful", but it is directly accessible to another user. A "flag" indicates simply that it is new and is therefore to be treated with caution. Experts are designated according to the class of data and the usage context. Periodically, they receive lists of recently-created records and, using special interfaces, they correct or validate these data according to the creation context. A validated item of data loses its control record and is said to be "certified".

- the management of the accesses to the database:

The accesses are subject to authorizations which depend on the project with which the user is concerned and his level of expertise. The access authorizations can be modified by the database administrator; they are managed in accordance with the following hierarchy:

- level 1: database
- level 2: specific project
 - a. collaborator
 - b. expert in a specific class of data
- level 3: common data.

A user who is authorized to access only the first level has access to the database within his own personal context and has tools which enable him to manage his gatherings (herbarium labels, gathering maps, lists) and his bibliographical data. A user who is authorized to access the second level may have access to the data of the project in which he is participating. If he is a collaborator, he can perform research and therefore read the data, acquire new data and modify them, provided that he is a creator of these and they have not yet been certified. If he is an expert in a domain which is specific to the project, he will have the task of validating the new data for this field.

The user who is authorized to use the third level has access to the data that is shared between different projects such as, for example, the abbreviations of authors' names. Such an expert is responsible for standardizing the information at institution level; he is the guarantor of the quality of this information.

The checklist context

The checklist of the Chaco represents a specific usage context of the general database. We shall not, at this point, go into the detail of the architecture of the database. A simplified entity-relationship diagram (CHEN, 1976) (Fig. 1) enables it to be referred to without entering into the specific problems of implementation. The classes of data, and the part they play in the project, are described below.

Geography:

The herbarium samples are the first information source for preparing the checklist. The gathering localities must be defined in a geographical reference system to enable appropriate exploitation of the data from the samples. A *locality* is a geographical point characterized by the following attributes: name_of_locality, latitude, longitude, altitude, type_of_locality, reference_map. Two distinct hierarchies coexist in the geographical field: one is political (*country*), the other is flora-related or phytogeographical (*region*). The political-based hierarchy contains between one and three levels: the country is always the first since, depending on the political breakdown of the country concerned, one or two subdivisions are included. Each *locality* is necessarily comprised in a political unit. A *locality* can be included in one or several *regions*.

The geographical context of the Chaco is defined in database terms according to the division chosen (Fig. 2). The user has access to a gazetteer assembling the *localities* comprised in the limits of the division.

Two strategies of data acquisition are possible at the level of the gazetteer: either data acquisition when new localities appear on the samples, or the preparation of an "a priori" gazetteer. A digitalizing table coupled with a PC enables the localities to be identified from a reference map; the co-ordinates are calculated by the computer. A set of conventions has been drawn up for identifying the objects from the map. This technique enables a basic gazetteer to be rapidly prepared.

Synonyms are easily created at the level of the *localities*, which makes it possible to manage data from different sources, for example in different languages.

Specimens:

This class of data is complex. The scheme can be read in the following manner: a *gathering* originates from a *locality*, is made by a *group* of one to many *persons* and comprises at least one *sample*, which carries zero to many *determinations*, performed by a *group* of one to many *persons*.

A *gathering* contains the data of the label: collector_number, date_of_collection, description_of_locality, station, substrate, etc. The *samples* of one and the same *gathering* can be distinguished from one another by the attributes: deposit_herbarium, typus. Among the *determinations* it is always possible to distinguish the name given on the label and the successive determinants. The *determinations* apply either to the entire gathering, or to a specific *sample*. The collectors and the determinators are treated as *groups* of *persons* (see below).

Within the framework of the Chaco checklist, the collections primarily processed are the Chaco herbaria present in Geneva. Other Chaco collections outside the CJB are also exploited. Furthermore, gatherings are envisaged for filling gaps in geographical coverage. A simplified

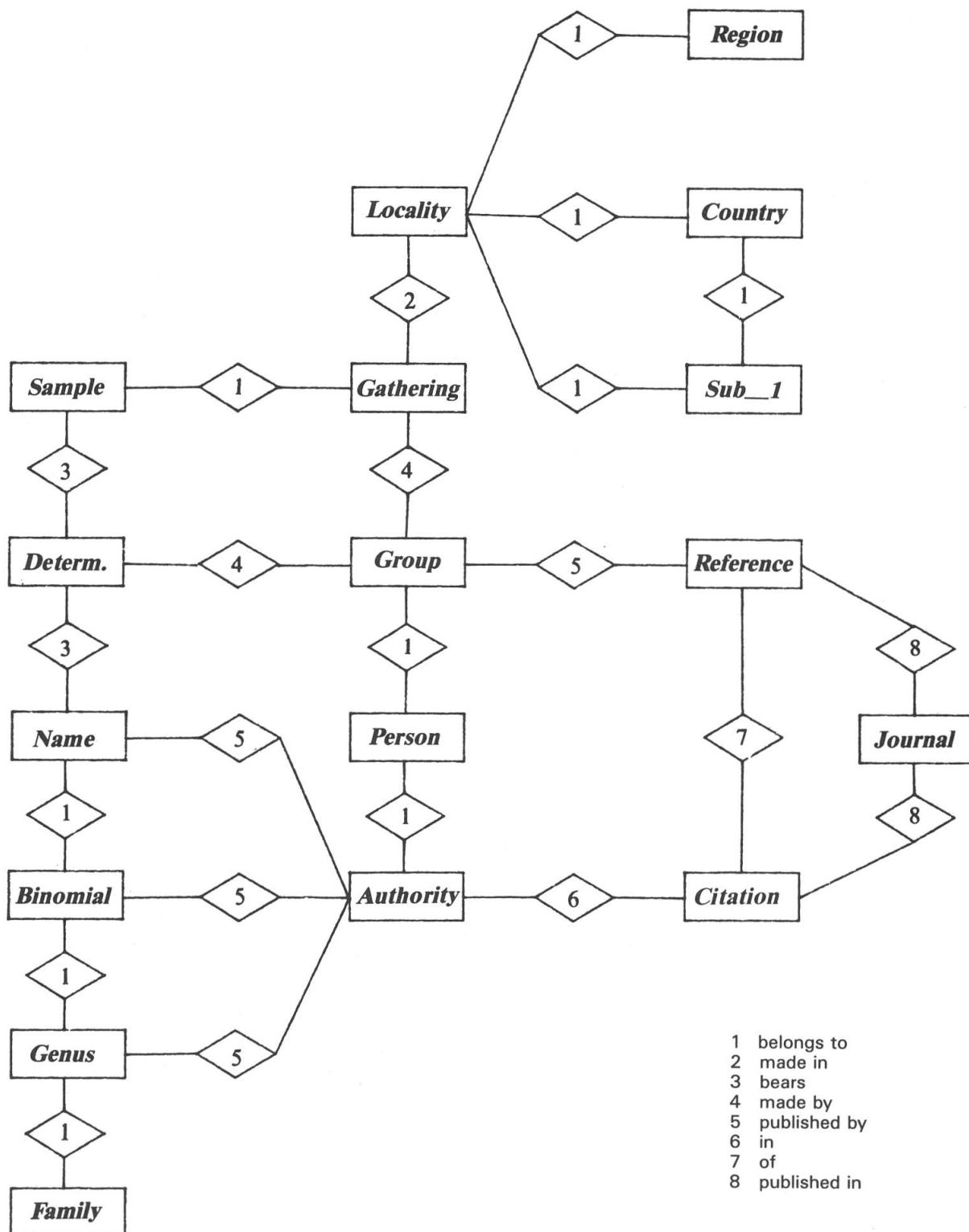
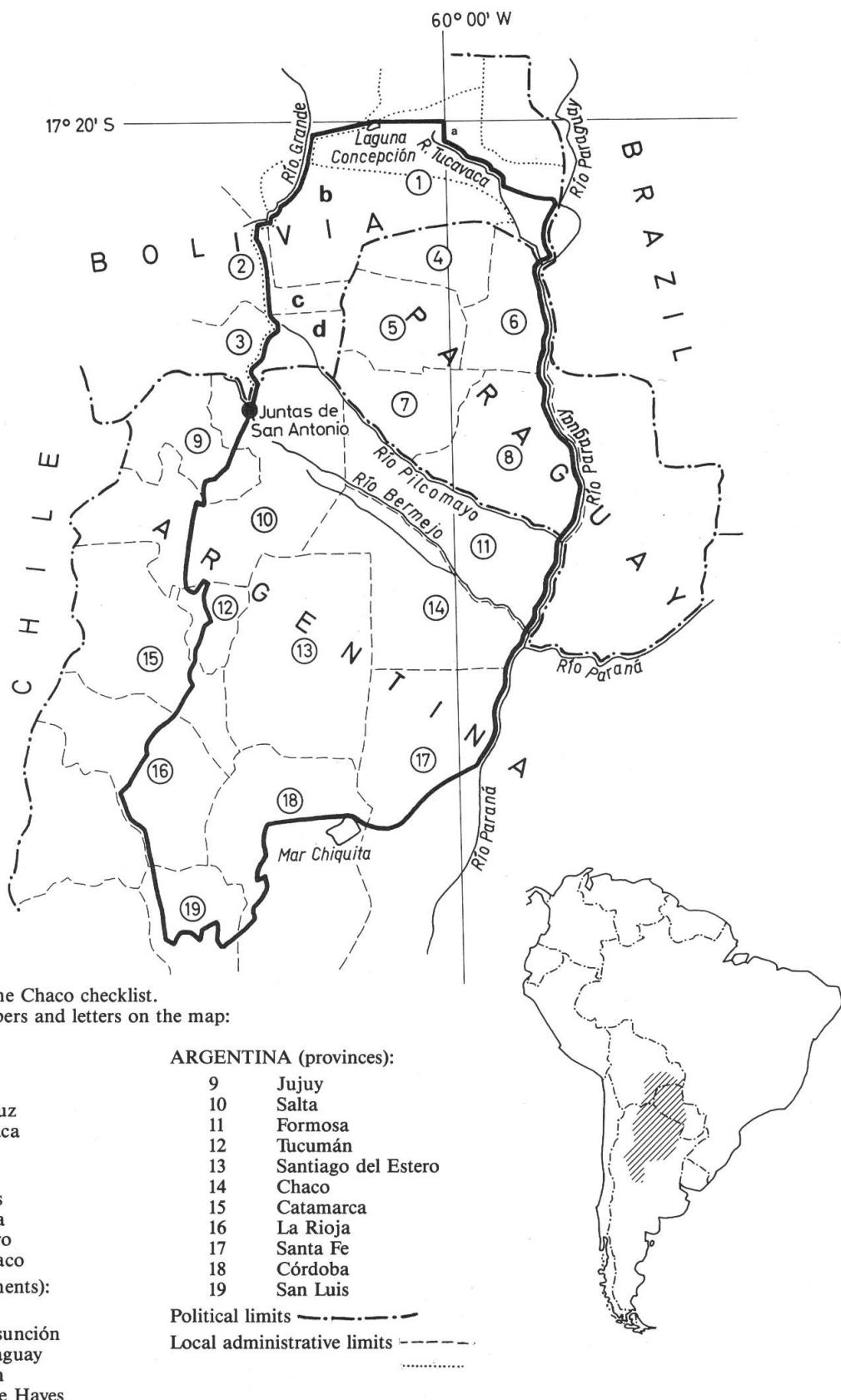


Fig. 1. — Simplified entity-relationship diagram of the checklist context of the integrated database system.



application thus enables the collectors of the CJB to acquire their gathering data and obtain herbarium labels. The link with a gazetteer selected by the user is effected interactively. If these gatherings concern the checklist project, they can be transferred directly into the *specimen* data and exploited within the framework of the project.

Persons:

This data covers the collectors, the determinators, the authors and name authorities. The notions of *group* and of *authority* enable several *persons* to be associated for a given task. A *group* comprises 1 to n *persons* who are collectors, determinators or authors; the name_of_person and initials_of_forenames are the attributes used. The case of citations of name authors is different. A simplification is applied, which makes it possible to limit to four the maximum number of names cited: the names preceding the particles *ex* are not recorded and in the case where more than two authors sign a contribution, the expression *& al.* follows the name of the first author. The relation *authority* combines between one and four abbreviations of author's names (or *& al.*).

In the particular case of *authorities*, the notion of usage context is important. Certain projects, in particular in the Mediterranean, adopt the abbreviations of author's names according to the MED-CHECKLIST (GREUTER & al., 1984-). The Chaco checklist will be prepared citing in so far as possible the abbreviations of names of authors according to the "Draft index of author abbreviations" (MEIKLE, 1984). In so far as possible, the two versions of abbreviations are recorded in the database in order to be able to cite an author according to either of these standards.

A *person* who is a name author, collector and determinator will be recorded only once in the database. The role of expert is indispensable at this level. His role is then not limited to the usage context of a project, but extends to the whole of the institution. The new data are periodically indicated to the expert. He is responsible for completing, if possible, the missing information, for assembling the redundant information and for validating the whole.

Nomenclature:

This class of data contains the elements necessary for the complete citation of a name. To be validated, a name must consist of at least the following data: name_of_genus (existing in the database, thus related to a family or even a sub-family), species_epithet and *authority*. The name of infraspecific rank comprise, in addition: infraspecific_rank, infraspecific_epithet and *authority*. In the context of the checklist, the name is completed by the citation of the protologue: title_of_periodical or title_of_work, volume_number, page_of_citation, year_of_publication, comments. The abbreviations of names of periodicals are treated in accordance with the "Botanico-Periodicum-Huntianum" and the titles of books follow the "TL-2" (STAFLEU & COWAN, 1976-1988).

The minimum hierarchy chosen as regards names is: family, genus, species, infraspecific rank, but other ranks can be added. The creation of a name at the rank of sub-species is possibly only after having searched for, or created in the database, the name of the species to which it belongs.

As for the *person* data, validation at overall level is effected by an expert.

The components required for the preparation of the preliminary checklist of the Chaco, as well as their processing in the integrated database, are briefly described above. As for the distribution of the checklist, several possibilities are available. Of course, the checklist can take the conventional form of a list of taxon names noted during the analysis of the collections. However, in view of the functions offered by the system, multiple selections are possible according to users' requirements. Thus the preliminary checklist designed as a working tool will fulfil its role in full.

Editor of the Chaco checklist

According to HUEK (1978), the Chaco occupies a vast alluvial plain located in the centre of South America. It extends approximately from Santa Cruz in Bolivia, in the north, to Mar Chiquita in Argentina, in the south; the tropic of Capricorn passes through its centre. To the west, the Chaco extends as far as the foothills of the Andes, while to the east, it slightly runs over the natural frontier

formed by the Río Paraguay and then the Río Paraná. It is therefore an extensive geographical zone, of about 1·000·000 km², covering the south-east of Bolivia, the western half of Paraguay, a small part of the south-east of Brazil and the quasi-totality of northern Argentina. It was around this large territory that a frontier delimiting the dition of the Chaco checklist had to be traced. Two possibilities are therefore open: either to select purely political limits (borders of countries, departments, provinces, etc) or to adopt a purely phyto-geographical standpoint, and in that case an attempt is made to adhere to a limit of Chaco vegetation in the strict sense of the term. There are therefore two opposing visions: one being artificial, the other natural. For reasons of facility and work efficiency, the solution adopted for the Chaco checklist is a mixture of these two points of view.

For Bolivia, there is no modern statement of the Chaco vegetation and, consequently, no precise representation of what the natural limits would be of this formation. The frontier therefore passes through points which can be identified on a map, respecting as far as possible the vision of HUEK (1978). The limit thus obtained is more artificial than natural.

As for Paraguay, the delimitation appears to be simpler, since the Chaco vegetation occupies the whole of the western part of the country. We thus arrive at an artificial (political) frontier which, grosso modo, follows the natural limits of the Paraguayan Chaco; the latter only slightly extends beyond the Río Paraguay.

Finally, for Argentina, we should note the existence of the "Flora Chaqueña" project (DIGILIO, 1971); this flora-related study covers the Provinces of Formosa, Chaco and Santiago del Estero, these three provinces being considered, by the authors of the project, as representative of the Argentinian Chaco flora. This is therefore a project whose authors have deliberately defined artificial limits. Nevertheless, the Chaco checklist must cover the overall territory of the Chaco, including in Argentina. This is why the natural limits of the "Parque Chaqueño", defined by RAGONESE (1959, cited in DIGILIO, 1971), have been adopted, but for two exceptions.

As indicated below, the dition of the Chaco checklist includes the foothills of the Andes. A limit of 1·500 m will however be adopted, since beyond this altitude the Chaco vegetation undergoes marked change or is even non-existent. Furthermore, this limit makes it possible to take into account the vegetation of the mountainous massifs present in the Chaco plain (Sierra de Santiago, Cerro Léon, Cerro Caimán, Cerro Cabrera, etc); a knowledge of this type of vegetation is of capital interest for an understanding of the Chaco system as a whole.

We should also note that the frontiers defined below are not absolute, but should rather be considered as transition zones. The frontiers of the dition are therefore thus defined (Fig. 2):

Bolivia (considered clockwise):

- west frontier of the Province of the Gran Chaco (Department of Tarija);
- west frontier of the Province of Luis Calvo (Department of Chuquisaca) up to the Río Grande;
- Río Grande as far as the Province of Chiquitos (Department of Santa Cruz); in view of the important impact of man-made factors on the vegetation of the region surrounding Santa Cruz (essentially agriculture), no account has been taken of the zone located to the west of the Río Grande;
- frontier of the Province of Chiquitos up to the Laguna Concepción;
- parallel 17°20'S as far as the meridian 60°00'W;
- meridian 60°00'W as far as the Río Tucavaca (intermittent part of the rio);
- Río Tucavaca as far as the railway linking Santa Cruz to Corumbá (as far as Tucavaca);
- railway as far as Puerto Suarez (frontier between Bolivia and Brazil);
- frontier between Bolivia and Brazil, as far as Paraguay.

Paraguay:

The dition of the Chaco includes the whole of western Paraguay, in other words, the territory located to the west of the Río Paraguay, which represents its eastern limit. The following Departments are therefore included: the Chaco, Nueva Asunción, Alto Paraguay, Boquerón and Presidente Hayes.

Argentina:

We shall adopt the limits of the Argentine Chaco, as defined by RAGONESE (1959, cited in DIGILIO, 1971).

These frontiers have been modified at two points:

- the north-west frontier has been moved further west, as far as Juntas de San Antonio, and rejoins the initial profile above San Salvador de Jujuy;
- the north-east frontier is defined by the Río Paraná.

Our dition therefore includes the following Provinces:

- in their entirety: Formosa, Chaco and Santiago del Estero;
- in part: Jujuy, Salta, Tucumán, Catamarca, La Rioja, Santa Fe, Córdoba and San Luis.

Conclusion

The preliminary checklist of the Chaco results from the study of the Geneva herbarium specimens. The Chaco collections of other botanical institutions in Europe and America will subsequently be included. This checklist is a working tool designed to develop in several directions. The integration of additional data, essentially bibliographical, will provide a complete information system on the Chaco. This procedure will, in particular, make it possible to present the checklist based on modern taxonomic views.

At the origin of the checklist of the Chaco is located the integrated database system set up at the CJB, designed from the outset as a system matching the overall scale of the institution. The checklist of the Chaco is a specific view of this general database. In particular it represents the Chaco window of the "Flora del Paraguay" and data can obviously be shared between the two projects. In this respect, it constitutes a source of information for the "Flora del Paraguay" collaborators. It must, however, be stressed that certain data is common to other floristic projects, as the "Flore de Corse" or Alpine flora studies, but also to systematic and bibliographic ones. This integrated database system, continually evolving through several development stages, might be considered as a significative improvement of the CJB abilities.

REFERENCES

ALLKIN, R. (1988). Taxonomically intelligent database programs. In: HAWKSWORTH, D. L. (ed.), *Prospects in Systematics*: 315-331. Clarendon Press, Oxford.

BISBY, F. A. (1984). Automated taxonomic information systems. In: HEYWOOD, V. H. & D. M. MOORE (ed.), *Current concepts in plant taxonomy*: 301-322. Academic Press, London.

BOCQUET, G. (1982). Activités des Conservatoires et Jardin botaniques en 1981. *Candollea* 37: v-xlii.

CHEN, P. P-S. (1976). The entity-relationship model: toward a unified view of data. *ACM Trans. Database Syst.* 1: 9-36.

CROSBY, M. R. & R. E. MAGILL (1989). *TROPICOS. A botanical database system at the Missouri Botanical Garden. "The booklet"*. Missouri Bot. Garden, St. Louis.

DALLWITZ, M. J. (1984). *User's guide to the DELTA system. A general system for coding taxonomic descriptions* ed. 2. CSIRO, Canberra.

DIGILIO, A. P. L. (1971). Proposito de estas notas. *Notas Prelim. Fl. Chaq.* 1: 1, fig. 1.

FROST-OLSEN, P. & L. B. HOLM-NIELSEN (1986). A brief introduction to the AAU-Flora of Ecuador information system. *Reports from the botanical institute, university of Aarhus* 14: ii, 1-39.

GREUTER, W., H. M. BURDET & G. LONG (1984-). *Med-Checklist. A critical inventory of vascular plants of the circum-mediterranean countries*. Vol. 1, 3, 4. CJB, Genève & Bot. Museum, Berlin.

HUECK, K. (1978). Los Bosques de Sudamérica: ecología, composición e importancia económica. Soc. Alemana de Coo. Tecn. (GTZ), Eschborn.

MASCHERPA, J.-M. & G. BOCQUET (1984). An outline for a database within a major herbarium. In: ALLKIN, R. & F. A. BISBY (ed.), *Databases in Systematics*: 235-248. Academic Press, London.

MEIKLE, R. D. (1984). *Draft index of author abbreviations compiled at the Herbarium Royal Botanic Gardens, Kew*. H. M. S. O.

RAMELLA, L. & R. SPICHIGER (1989). Interpretación preliminar del medio físico y de la vegetación del Chaco boreal. Contribución al estudio de la flora y de la vegetación del Chaco. I. *Candollea* 44: 639-680.

SKOV, F. (1989). Hypertaxonomy — A new computer tool for revisional work. *Taxon* 38: 582-590.

SPICHIGER, R. & J.-M. MASCHERPA (1983). Guía para los autores. In: SPICHIGER, R. & G. BOCQUET (ed.), *Flora del Paraguay. Serie especial 1*. CJB, Genève & Missouri Bot. Garden, St-Louis.

STAFLEU, F. A. & R. S. COWAN (1976-1988). *Taxonomic literature* ed. 2. 7 vol. Bohm, Scheltema & Holkema, Utrecht-Antwerpen & Junk, The Hague-Boston.