ALICE software news

ISSN 1369-4219

SIX

contents

introduction

Potatoes and palms are on the menu for readers of this, the sixth, issue of Alice Software News (ASN). Several articles describe larger collaborative projects and a staple diet is provided through articles describing developments of both Alice itself and of third party programs.

Readers might like to note that ASN is now registered with the International Standard Serial Number: ISSN 1369-4219. This registration no. will be useful for your librarian and should be used when referring to articles included in this and future issues.

So, what's in this issue?

A theme common to many contributions to this issue is the management and dissemination of information about economically important plants. Articles describe the use of Alice in a Rattan palm database, in a Brazilian regional Plant information service which links academics (with knowledge about plants) to local farmers (that use them), a project in Darwin (where else?), Australia, which aims to conserve aboriginal knowledge about plants from the 'bush' and a new European Union funded initiative which is managing descriptions of potato germplasm collections in a distributed, collaborative fashion. Some Alice users will have already heard of the Species 2000 and IOPI Checklist projects: others will not. The organisers of these collaborative, international projects describe their aims and how Alice users may participate.

Technical descriptions of the Alice System, meanwhile, highlight important improvements to data quality in Systems 2.1 and 3.0, a description of System 3.0, and details of new programs appearing within Alice itself and as third-party tools. We announce a new program to let you manage images within your Alice database - so start digitising now! Technical notes describe what sort of computer you should look for when building Alice System 2.1 databases and, sorry, not even we could resist it, there's a mention of the millennium bug and its significance to Alice users.

Do let us know what you would like to hear about in the next issue.

The Editors

Improvements to *Ace*

Ace is the new data entry and data editor program in Systems 2.1 and 3.0. It replaces the *Alice*, *Atext* and *Nview* programs found in System 2.0. When we first released *Ace,* it already did much more than all of these older programs put together (see last issue of ASN). Since we released System 2.1, however, many new features have been added to *Ace* and are available in both Systems. Some of these improvements are listed here (see below).

Adding so many new functions to *Ace* made it memory hungry. Various users also requested that we separate more clearly:

- tasks involved in setting up and designing a database (carried out by the project leader)
- tasks associated with entering data (commonly carried out by project staff, students etc.).

These parallel pressures led to the subdivision of *Ace* into two programs: *Ace_designer* and *Ace_builder.* We can draw a parallel between building databases and building houses - the former is a tool for use by the architect while the latter is for use by carpenters, bricklayers and plumbers.

Improvements to *Ace* since release in 1997

SYSTEM-WIDE

Users can
- view reports generated by *Ace* using the 'F4' key from any menu
- import and edit the text for database credits
- set prompts 'on' or 'off' to confirm data entry. These case be set for each datatype independently.
- set 'on' or 'off', again for each datatype independently, prompts to remind you to attach citations to each new data entry.

Ace
- marks databases as 'unhealthy' when operations which involve entering or changing data do not complete successfully ('transaction monitoring' - see article on data quality and security).
- does not allow you to open databases previously marked 'unhealthy'.

NAMES & TAXA

Users can
- count/view/write those scientific names which contain a particular name element (authority or epithet).
- assign taxa to an <unknown> higher taxon.
- view either ALL names retrieved so far in a search, or ONLY those names of the current rank. This makes it easier to select taxa which have no subtaxa.
- select and work with an existing homonym as well as reviewing its synonymy and citations.
- select taxa directly from Recall Taxon menu.
- display taxonomic ranks in hierarchical order in all browse windows.

Ace
- uses an improved name sorting algorithm.
- has enhanced defences against the unexpected or illegal creation of homonyms when editing name dictionaries.

DATA SOURCES

Users can
- count how many times an individual data source is cited and where.
- access the data source editor whilst entering data for taxa.
- sort and select data sources on a combined index of author+title as well as on the title etc. alone.

DESCRIPTORS & TAXON DESCRIPTIONS

Users can
- reorder and insert states in the definitions of ordered descriptors.
- obtain improved descriptor statistics reports.
- count/view/write taxon records for a particular descriptor or state
- count/view/write taxon records for a particular value for other data classes (e.g. a place or a use).

TEXT INFORMATION

Ace
- generates structured notes automatically for ALL operations that change names or taxa (e.g. add synonym, edit name, change taxon status, delete name, change preferred name etc.).
- supports foreign languages for notes.
- allows users to select default text editors from menus.
- includes free text descriptions within taxon reports.

Images
in Alice databases

AliceWeb users can already include a number of images of each taxon in the HTML pages that AliceWeb generates from an Alice database. Eduardo Dalcin (see ASN issue 5) developed a special tool to associate images with particular taxa, for this purpose. This works well, but maintaining these taxon / image links can be a frustrating experience for those with lots of images and / or taxa. It would be far simpler to store and manage them within your Alice database. AliceWeb and the Alice System are being modified to facilitate management and publication of images.

A NEW PROGRAM

We are developing a new program, as part of the Alice System, which will manage images within Alice databases. The program, planned for release in the last quarter of 1998, will be the first to use the new Alice graphical user interface (GUI) being developed for the whole System. This first release of the program will let you:

- find and view selected images.
- look up and select taxa using their preferred names.
- associate an image with a selected taxon.
- view or delete images for a selected taxon.

We have yet to decide upon the name for this program. Any suggestions (well almost any) are welcome and we will announce its name in the next issue of ASN.

HOW CAN YOU PREPARE?

If you wish to use images in your Alice database and publications you can now begin the task of collating, digitising, sorting and documenting the images you wish to include. You can store your images in any one of the standard image formats (*.jpg, *.gif for example).

FUTURE DEVELOPMENTS

Subsequent releases of this new program will provide:

- full nomenclatural lookup with synonym cross referencing
- an illustrated glossary - enabling images to be attached to descriptor state definitions (e.g. 'obovate' or 'pedunculate')
- illustrated dictionaries - enabling images to be attached to economic importance, or habitats etc. (e.g. 'thatched roof')
- illustrated taxon descriptions - enabling images to be attached to individual observations of taxa ('leaf shape for species X' or 'Wing pattern for species Y').

BEYOND IMAGES

A very similar approach can be taken for the storage of other kinds of information (for example sounds, e.g. bird song, documents and web links) and we would like to hear from any readers that are interested in storing such kinds of information.

NEW USER GUIDE AVAILABLE FOR **ACE**

User-Guides for *Ace* now exist for use with both Systems 2.1 and 3.0. The guides document *Ace* as it was released last year and we are treating this as a draft publication. We will incorporate in the final publication additions necessary to reflect the development of *Ace* and changes in response to users' reactions.

The Guide is large! 18 chapters are presented in two volumes in the same format (A5) as was used in our previous Guides. It is available, free of charge, to all registered users of System 2.1 and 3.0. Copies are also available electronically for anyone prepared to give us feedback to help improve the final version.

Improving data quality and s

QUALITY ASSURANCE AT DATA ENTRY

Ace has been more rigorously tested than all previous program releases. All known bugs in System 2.0 have been fixed. *Ace* incorporates more data validation checks than its predecessor in System 2.0.

ALICE ON NETWORKS

Alice Systems 2.1/3.0 can be used with 100% confidence to access databases held on networked server. Once opened by an Alice program, a database is automatically locked to protect against simultaneous, and potentially damaging, access by others. Anyone attempting to access a locked database is told that the database is in use and 'who' is using it.

ALICE IN MULTI-TASKING ENVIRONMENTS

More than one Alice program can also be run safely in multi-tasking environments like Windows 3.1, Windows 95/NT, and OS/2.

'TRANSACTION MONITORING'

Transaction monitoring has been implemented with the release of Alice System 2.1. A transaction is any attempt to add or modify information in an Alice database. Transactions can fail for many reasons, including power, hardware or software failure or even in changes to the way a computer is configured. Failure to complete a transaction can compromise database integrity. Alice System 2.1 programs detect such failures and, when they occur, prevent users from accessing a database until a full check of its health has been carried out using *Amie*.

Transaction monitoring has several benefits for users. Particularly important is the protection it affords of a user's investment in their data. Since potential problems are detected early, enforced use of *Amie* can prevent the creation of secondary errors that might occur if a user attempted to modify or add data to an unhealthy database.

Data quality has always been one of the Alice System's major strengths. Systems 2.1 and 3.0 have significant new features and new programs which reinforce this tradition, further improving the reliability and long-term security of your data. Both Systems are built upon a more sophisticated data structure which supports these improvements in data integrity as well as many other new features.

'DATABASE HEALTH'

Amie first checks that all rules for relational databases are properly followed, including referential integrity, then checks that all logical and biological rules that must be followed in Alice databases are properly applied. To be healthy, a database must satisfy both sets of rules.

A COMPREHENSIVE REWRITE OF AMIE FOR SYSTEM 2.1

Amie is used to check database health, reporting and, where possible, fixing errors, to warn of potential problems, to re-index databases and to physically remove unused or previously deleted records. In the right hands, *Amie for System 2.0* is a highly flexible tool for testing and fixing problems with either an entire Alice 2.0 database or for particular data types. Users can decide what tests are carried out and the order in which they are done. Many users, however, found this flexibility confusing and expressed the desire for an easier-to-use program.

Amie for System 2.1 is simpler in use than its predecessor. The new program requires little user interaction and has far fewer menu options. This version undertakes even more data validation tests than before. As before, *Amie* tests for, reports and then fixes errors. Wherever practical, the fixing of errors is automated.

A health check is invoked automatically if another application has marked a database as unhealthy.

In exceptional circumstances *Amie* may not be able to fix all the errors it detects. In these cases it may be possible to return the database to Alice Software for repair .

AMIE FOR SYSTEM 2.1 HAS:

- *improved database integrity testing;*
- *data recovery in the event of a failure during a session with Amie (see below);*
- *improved reporting;*
- *data compression options: 'pack', 'trim' and 'zap' for each data type independently.*

  a) 'pack' physically removes data records deleted by an Ace user;

  b) 'trim' removes all those entries in a particular Alice dictionary which have not been recorded for any taxa at all;

  c) 'zap' deletes all the entries in a particular dictionary and all records in its associated data

ecurity

tables regardless of whether data have been previously recorded for taxa or not;

A future release of Amie for System 2.1 *will allow users to create and manage databases archives.*

HOW WILL YOU USE AMIE?

The new functionality and simplified interface means that you will use *Amie* in a different way. Instead of using *Amie*, from time to time and only to report on one particular aspect of your data, you will now typically use *Amie* overnight regularly to automatically complete a health check, run all maintenance tasks and to generate a report of what has been found and what has been done.

DATABASE RECOVERY

Amie implements a new feature, 'database recovery', to guarantee against data loss. Each time that *Amie* begins work with a database, it automatically makes a copy of the original data. If, for any reason, a maintenance task fails, *Amie* recovers the saved copy of your database.

A NEW VERSION OF DBUPDATE

Dbupdate converts an Alice 2.0 database for use with Systems 2.1 or 3.0 and System 2.1 databases for use with System 3.0. *Dbupdate* does not reverse this process: it cannot convert System 3.0 databases to System 2.1 for example. To prevent the inheritance of errors from older databases, the new version of *Dbupdate,* by labelling newly converted databases as 'unhealthy', forces a heath check to be done by *Amie*.

System 2.1: releases

As announced in the last issue of ASN, formal release and controlled distribution of Alice System 2.1 began in 1997. In this issue of ASN, based upon the experience of early users of the new system, we propose the computer and operating system requirements for running System 2.1 (see page xxx).

Users migrating from System 2.0 have had varied experiences upgrading their System 2.0 databases. Some databases only required processing using *Dbupdate*, others required further processing using *Amie* reflecting the more exacting data integrity demands of System 2.1. These improvements are described in more detail on page xxx.

For advice on how to upgrade your database contact Alice Software.

We have significantly improved the new data entry and data editor program, *Ace*, has been upgraded (see page xxx). With release 8 *Ace* has been into two separate programs: *Ace_builder* and *Ace_designer*.

System 2.1 releases

Release	Description	Date
1.	First test version for use by Alice Software	Oct 1996
2.	Second test version for use by Alice Software	Dec 1996
3.	Single user version of System 2.1 for users.	Apr 1997
4.	Multi-user version with transaction monitoring.	Jun 1997
5.	Improved text management and descriptor editor.	Jul 1997
6.	Maintenance release	Sep 1997
7.	Improved query engine and query history manager for all programs. New version of name management module in *Ace.*	Oct 1997
8.	*Ace.*split into "database design" and "data entry" programs. First release of *Amie.*	April 1998
9.	New report formats in *Ace.*and *Awrite* with improved flexibility and more consistent listing of scientific names. Second release of *Amie.*	July 1998
10.	Maintenance release. Year 2000 compliant version of system 2.1/3.0. Third *Amie* release.	Dec. 1998

Alice System 3.0:
Progress and overview

What are the differences between Alice System 2.0 and Alice System 3.0? The answer is many if you work with infraspecific taxa.

Alice System 3.0 supports:

- Nine rather than four taxonomic ranks
- Informal ranks, such as 'cultivar' or 'cultivar group'
- Accessions or specimens below the 'lowest' taxonomic rank
- Two new classes of taxon: 'informal' and 'unidentified'
- alternative flavours of taxonomic system. It can be modified to allow for different conventions for taxonomic ranks. *Brassica* crops for example, use 'convarietas' and 'varietas' while in *Allium* 'cultivar group' is used instead.

System 3.0 is designed to be of use to those who wish to build databases containing descriptive information about infraspecific taxa. It is particularly useful, therefore, to those storing descriptions of cultivated plants, such as cultivar registration authorities, gene banks, horticulturists, plant breeders and garden plant specialists.

System 3.0 supports the more rigorous demands of the formal zoological and botanical nomenclatural codes at the upper taxonomic ranks while being more relaxed about the names used for 'informal' taxa at lower ranks, typically at the cultivar level. You can work with taxa of formal or informal rank together in one database and can switch quite easily between them. Data entry can be adjusted to avoid repetition when entering names (repeated entry of the species name in a database of crop cultivars for example).

System 3.0 allows accessions/specimens to be entered and associated with existing taxa in the database or classed as 'unidentified'. Descriptive data can be entered for accessions.

These developments required a considerable revision in the underlying Alice data model and changes to the interfaces for generating queries, writing reports and exporting data. The latter opens up the exciting potential for users of the system to take advantage of other tools, such as online identification and automatic key generation.

Two significant omissions from Alice System 3.0 currently are the *Sam tool kit*, for data import and for combining databases, and the *AliceWeb* hypertext publication tool.

Development of Alice System 3.0 arose from a particular request from a collaborative project between the cultivar registration section of the Scottish Agricultural Science Agency based in Edinburgh and the vegetable gene bank at Horticulture Research International, in Wellesbourne, England. Their purpose was to produce a database of descriptions for cultivars and accessions of *Allium*. The potential for using the modified system to manage descriptions for other species was recognised early during the life of the project. System 3.0 is now being used to store official descriptions for potatoes in a European Potato Cultivar Database project (see separate article) and a celery database.

Merging databases and importing data

Sam was the program used to merge System 2.0 databases and import data into a System 2.0 database. The program incorporated all of the necessary functionality within one program so that a database administrator need use only one program. Experience proved, however, that there were several disadvantages to this approach:

- Database administrators were less aware of the actual steps involved and their consequences;
- The database administratorÕs had less opportunity to intercede, influence the process or control reporting levels;
- Only more technically minded users felt comfortable using *Sam*;
- The functionality was not available for more general use;

In System 2.1 a *Sam tool kit* replaces the single program in System 2.0. The kit, which is still being developed, consists of a number of smaller tools. The complete tool kit still offers the necessary functionality for importing data or merging databases. Database administratorÕs can take a more flexible approach to the processing sequence for database merging, using the new tool kit. The tools are also of use to many Alice users, not just those interested in merging databases. For example, one tool, the database comparison engine, detects and reports differences between any two Alice databases.

A fuller description of the tool kit will be featured in a future issue of Alice Software News.

Global Plant Checklist

The diverse users of plants and their products around the world need a reliable, unified inventory of plant species. IOPI (the International Organization for Plant Information) aims to provide this with the Global Plant Checklist, initially covering the estimated 300,000 species of vascular plants.

The Checklist is a partner in the Species 2000 project, which aims to provide a central access point to all such global species databases. In addition, the Checklist is working closely with the Species Plantarum (SP): World Flora Project of IOPI, so that the names and distributional data gathered by contributors to the SP will be included in the Checklist, and links will be provided from the Checklist to the species' treatments in SP.

The full Checklist database will have a distributed structure, so that data on plant families will be held at different sites around the world, and edited and maintained there by the relevant group of botanical specialists. These elements of distribution and specialist-editing are central to the IOPI Checklist concept. However, funding limitations have not allowed us to implement those elements yet.

A more limited, provisional Checklist has been on-line for over a year, through the IOPI home page at http://iopi.csu.edu.au/iopi/. This is the first implementation of the `potential taxon concept', which is central to the IOPI data model - discussed by Walter Berendsohn in Taxon 44: 207-212 (1995) and Taxon 46: 283-309 (1997). This is a ground-breaking attempt to deal with all the complexities of botanical nomenclature. It is particularly important for names that have been misapplied, so that, for example, a species name `sensu Joe Bloggs' or `sensu Flora of Brobdignag', together with any associated data, can be assigned to the relevant taxon in the Checklist. This will be invaluable in sorting out common problems such as confused geographic distributions for taxa whose names have been misapplied or where a single species has been split into two or more taxa.

The provisional Checklist is already proving useful, both to enquirers and as a demonstration of the Checklist concept to potential sponsors, despite being incomplete in its world coverage and with limited botanical editing so far. We are adding existing electronic datasets (both regional and taxonomically based in scope), and aim to edit them botanically to unify all information for a particular taxon. Shortage of funding means progress is slow, particularly with implementing botanical editing of data. So far, regional datasets have been added for Australia, Europe, North America north of Mexico, and Peru, and family datasets for Casuarinaceae and Magnoliaceae. Various other datasets await incorporation.

Data in the Checklist is classified as being at one of three levels, depending on the extent of botanical editing by the relevant botanical coordinators: (1) Source datasets (the data shown as received from the dataset custodian); (2) Partly coordinated/edited; and (3) Fully coordinated. For an example of an edited entry in the provisional Checklist, look at *Casuarina equisetifolia*. However, until a data-editing tool has been written, botanical editing will proceed slowly. We expect this editing tool in 1998.

Future plans include

- finding funds to extend this provisional Checklist
- making the Checklist a distributed database with live links to original datasets, such as the ILDIS legume database
- developing a user-access module for easy editing of Checklist data by botanists
- providing alternative classifications for groups whose status or composition are not yet clarified
- links to other plant data, such as descriptions, illustrations, chemical constituents, regional treatments.
- finding funds to help holders of non-electronic datasets convert them to electronic form.

ALICE-users can contribute to this co-operative effort by the botanical community by making their plant datasets available, whether these are geographically or taxonomically based. Currently this entails sending the data for incorporation in the provisional Checklist. The suggested format for doing this is given on the Checklist home page. Once the Checklist is fully operational, live links will be possible to those ALICE databases that are connected to the Web.

This project has had input from more than 40 individual or institutional members in 27 countries. It depends on its members for success, and I invite ALICE-users to swell those numbers.

Karen Wilson
Convener of IOPI Global Plant Checklist Committee
Royal Botanic Gardens,
Mrs Macquaries Road SYDNEY
NSW 2000 AUSTRALIA
karen@rbgsyd.gov.au

Conservation of PotatoGenetic Resources in E

The potato (Solanum tuberosum L.) is the most important root crop of cool temperate regions and on a world scale, ranks as the fourth major crop after wheat, rice and maize. The genetic variation of the species is narrowly based. To exploit this resource fully, researchers require access not only to the germplasm but also all the information associated with it.

THE CONSERVATION, CHARACTERISATION AND USE OF POTATO CULTIVARS

The primary purpose of a European Union funded project on "The conservation, characterisation and utilisation of secondary potato cultivars for ecological production" is to increase the knowledge about what germplasm is available, and to encourage the use of this information by making it available to plant breeders and others. The main project comprises two sub-projects, the first concentrating on species of *Solanum* the second on cultivars and breeder's lines of *Solanum tuberosum* L. Twenty-five institutes from fifteen European countries are collaborating in the second sub-project, which started in 1996 and is due to be complete in 2000. Each participating institute is supplying descriptions of the potato germ plasm held in their collections. These are being collated in an Alice System 3.0 database and a copy of this central database is being returned to each participant. The collation, verification, presentation and dissemination of this information are being carried out by the Scottish Agricultural Science Agency (SASA), based in Edinburgh, Scotland. It has been agreed that SASA will generate a combined database that will be distributed to all the participants and will also make this information available on the World-wide Web.

Up to ninety-seven descriptors are recorded for each cultivar or breeder's line. These include its name(s), passport information (including country of origin, parentage and whether a cultivar has been awarded Plant Breeder's Rights), morphology, resistance to 33 pests and diseases, crop yield factors, horticultural properties and cooking qualities. The amount of descriptive data supplied by the institutes varies widely. Some are providing only passport data, others are donating large, comprehensive data sets. Some of the cultivars are being 'characterised' during the lifetime of the project. This involves growing plants in field trials and recording descriptions for them. As new characterisation data become available the enlarged data sets are forwarded to the co-ordinating centre again.

The first data sets were supplied in an electronic format during 1997. These have been translated into a simple exchange format and entered into an Alice database at SASA.

THE USE OF ALICE SYSTEM 3.0

The combined database is being distributed to all participants with the following Alice System programs:

- *Aview*, which will allow them to view and print individual cultivar descriptions and to produce nomenclatural reports, and;
- *Aquery*, which can be used to list those cultivars with any chosen set of properties.

A combination of the *Alex* program and *AliceWeb* for System 2.1 have been used to produce a mock set of web pages for display and dissemination on the WWW from the current version of the main database. A sample set of pages was recently demonstrated at an international project progress meeting held in March 1998 in Brest, France. A version of *AliceWeb*, modified for use with Alice System 3.0, is planned to accommodate the additional types of nomenclatural data stored within Alice 3.0 databases.

DATA PROCESSING

Combining data sets from different sources, which may represent different opinions, into a single database poses a number of challenges. The degree of agreement between descriptions of the same genetic material from different institutes, for example, can vary. This may be because either there are genuine location effects or the material is genetically heterogeneous or institutes are actually dealing with different cultivars. Where there is consensus that the differences in description are due to location, descriptors, that would be treated as 'not variable', i.e. should only have a single state, are defined as 'variable'. Each data observation in the database is cited using the name of the institute from which it comes. This makes it possible for the SASA team and users of the combined database to see where a particular observation was recorded.

A more significant problem for those donating data sets and those merging them will be deciding what constitutes a cultivar. These difficulties arise for two reasons; firstly because the descriptions differ so dramatically that it seems improbably that the two institutes are handling the same material or secondly because naming conventions differ. A simple example illustrates the problem. The cultivar Arnika occurs with 4 different spellings (Arnica, Arnika, Arnika (1914), and Arnika (1988) and with three different sets of parentage (Bintje x SVP 51-269, Granola * 71.241/50 and

Europe

Deutches Reich * Jubel). Such differences in the spelling of names and the use of synonyms need to be resolved before the final combined database is produced. An early estimate is that the database will include approximately 4,000 cultivars and breeder's lines.

Alice System 3.0, unlike its predecessors, does not yet have a version of the Sam tool kit for automating database merges. A restrictive and simplified approach is therefore being taken to building the central database. A simplified approach to merging is possible because the cultivars or breeder's lines belong to a single species. Such an approach would not be possible where there were complex relationships between names and taxa, as, for example, would be the case when merging species-based databases. Despite this simplification, merging databases, as described below, is restrictive. No editing is allowed on the central database, although SASA or others using Alice do edit their own institutional data sets regularly. Instead, the central data set is thrown away periodically as revised data sets are received. A restricted approach is also being taken to the detection and prevention of data conflicts when the data sets are appended to the central database. This leads to data quality problems that have to be resolved later. While not ideal, this approach has allowed SASA to continue to bring diverse sets of information together for comparison and dissemination.

The combined database is produced in four steps. Step 1 is to create a combined flat table. Step 2 is to generate data checking reports from this table. Step 3 involves comparing and validating the data against a reference set of valid states for each descriptor to produce a check report for use by crop experts. The final step uses a tool kit, built for our purposes, to create a new Alice database from the validated table. This is achieved by firstly generating the taxa in the database with their names and then appending the descriptors and citations to that database. Structured notes, another Alice data type, are then generated from the parentage information and accompanying remarks and appended to the database

A VALUABLE RESOURCE

This cultivar database is already being recognised as a valuable resource of interest to scientists working with germ plasm collections and to plant breeders interested in selecting material for plant breeding programmes. The database will be easy to distribute in an variety of attractive and easily accessible ways which can be generated automatically. Use of Alice will make it easy to identify material quickly with any desired combination of characteristics. Future plans include enriching the data content and incorporating images.

For further information about the project please contact the authors.

R. Grant D. Turnbull, P. Winfield
Scottish Agricultural Science Agency
East Craigs
Edinburgh EH12 8NJ
Scotland
Grant@sasa.gov.uk

A 'read-only' version of Alice

Database projects wishing to 'publish' their data electronically have sought a 'read-only' version of the Alice System to enable them to distribute some, or all of their data to other people. The End-user's Edition of Systems 2.1 and 3.0 can be distributed with data subsets. It allows those receiving the data to explore them and make databases searches without being able to make any changes to the data.

The key applications provided with the End-users' edition are *Aquery* - the Alice general purpose query engine - and *Aview*.

Aquery is very familiar to existing Alice users. It will allow searches of the database to be designed using any data item (place, habitat, use, name or any descriptor) using search operators ('Equal', 'Not Equal', 'Missing' and 'Unique'). Ordered descriptors can also be searched using comparative operators such as 'Less than' or 'Greater than or equal to'. Individual search elements can be combined using logical operators 'And' and 'Or'. Complex searches and their results can be stored for future use using a Query History Manager which also allows you to edit, rename or delete previous queries.

Aview allows you to

1. print simple pre-defined reports or checklists.
2. view the database statistics.
3. make nomenclatural searches.
4. find and select a taxon using any of its names (common names, preferred or synonyms).
5. view all or any subset of the data recorded for a chosen taxon.
6. discover the data source(s) of any individual data record.
7. ask, for each descriptor state or other data category used in the database:
   a) How many taxa are recorded for this descriptor / descriptor state / data category?
   b) Which taxa are recorded for this descriptor / descriptor state / data category?
   c) Print a list of taxa recorded for this descriptor / descriptor state / data category?

Plantas do Nordeste

Plantas do Nordeste (PNE) is an Anglo-Brazilian collaborative and interdisciplinary programme involving both research and information dissemination in the semi-arid Northeastern region of Brazil. The programme, which has been running since 1992, aims to promote the sustainable use of the plants of the region for the benefit of local communities. Our guiding theme is "local plants for local people".

Twelve projects have been funded since 1992, from a variety of sources, and have broadly fallen into two categories: "Biodiversity" (primarily concerned with the study and conservation of the flora in priority vegetation types) and "Economic botany" (concerned primarily with the sustainable use and management of forests / plants for the benefit of local communities, e.g. to provide better quality sources of goat forage throughout the year or affordable sources of medicines). Individual projects have been undertaken by Brazilian scientists, often in collaboration with partners from the Royal Botanic Gardens, Kew (UK). Most projects have direct links with one or more community groups. Each has a data management and dissemination component and is supported by a separate "Information, Dissemination and Training subprogramme" (see separate box) which provides mechanisms by which PNE's results can be disseminated more effectively throughout the region.

More than thirty Brazilian organizations participate in Plantas do Nordeste: Federal or State research agencies, universities, non-governmental organizations (NGOs) and grassroots organisations involved in alternative agriculture, forestry and community development. The PNE programme is co-ordinated by an NGO in Recife, Pernambuco with a Board and a small executive office. For further information about the PNE programme please contact:

Dr Jair Virginio, Associação PNE, Av. Gen. San Martim 1371IPA – Bloco 7 – Bongi, Recife, PE, Brazil 50761-000.
Email: jair@elogica.com.br

Northeast Brazil

- **Northeast Brazil consists of 9 states (from Bahia in the south to Maranhão in the north) and covers 1.5 million square km (about six times the size of the UK). It is home to c. 44 million people (30% of the Brazilian population).**
- **The region is a major centre of biodiversity and has possibly the greatest variety of ecosystems of any in Brazil. There is still much to be learnt about the flora. We know neither how many species are present (c. 20,000) nor what they are.**
- **The region is characterized by social and environmental problems resulting from a natural ten year drought cycle and decades of unsustainable exploitation of the natural resources. 22 million people live in poverty, of whom more than half live in rural areas.**
- **Rural poverty has led to large-scale migrations to major cities (particularly in the south of Brazil) and to the Amazon, which exacerbates social and environmental problems in those places.**

MAJOR NEW FUNDING FOR PNE'S INFORMATION DISSEMINATION PROGRAMME

Plantas do Nordeste (PNE) is taking an innovative approach to the organisation, exchange and delivery of information about plant biodiversity. Scientists that generate new information are working in partnership with those dealing directly with local communities and striving to benefit them through alternative and more sustainable use of their plant resources.

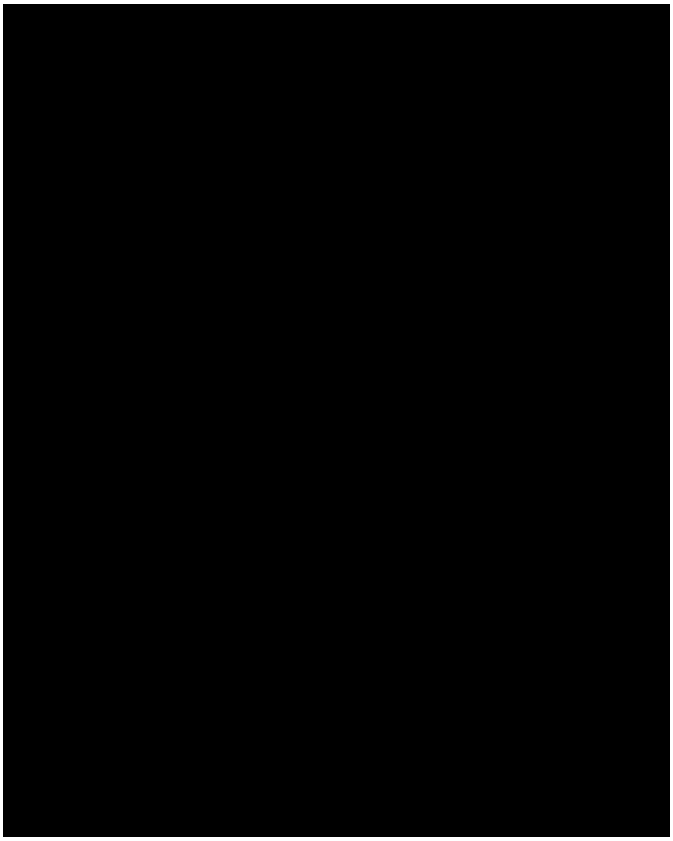
In 1997, the UK Government's Department for International Development (formerly ODA) approved a major funding (£1.85 million) towards PNE's Subprogramme for Information, Dissemination and Training (SIDT) during the next five years (1998 - 2002). Three units are being created to implement the programme:

1. a plant information centre (CNIP) in the Federal University of Pernambuco, Recife.
2. a community extension unit in the Recife Headquarters of AS-PTA - a Brazilian national development NGO.
3. an information repatriation unit in the Royal Botanic Gardens Kew, London.

PNE gratefully acknowledges major support from the Weston Foundation (UK) and from CNPq (Brasília). The latter provided awards enabling training courses and more than 100 studentships. Generous support is also acknowledged from the Darwin Initiative, the Reuters Foundation, RBG Kew, the Rufford Foundation, Shell (UK), Souza Cruz (Brazil) and several individuals.

The **plant information centre** will:

- offer advice to PNE projects and partners about biodiversity data management;
- create a regional documentation centre serving scientists, foresters and agronomists;
- collate and co-ordinate the results from PNE's projects together with existing data from the literature, collections and existing databases;
- disseminate information of practical value to those working with local communities;
- provide a technical information service to scientists and those agencies responsible for planning, development and conservation;
- train Brazilian scientists in the delivery and management of information.



The **extension unit** will:

- coordinate dissemination to PNE's priority target users in poor rural communities and intermediary agencies who work with them;
- establish pilot projects to work directly with farmers to promote and assess the farmers' use of PNE's information;
- hold user needs and feedback workshops, to ensure the demand-led nature of the service and to multiply the benefits of PNE throughout the region;
- produce a non-technical newsletter to promote lateral communication between users;

The **information repatriation** unit will:

- repatriate information about plants of Northeast Brazil from Kew Gardens and other European institutes. This includes electronically stored data and images, botanical expertise and text information stored in specialist libraries and collections;
- link relevant expertise at Kew and other European Institutions to the project in Brazil;
- support Brazilian visitors and staff at Kew that work with data about the plants of Northeast Brazil.

For further information about PNE's Subprogramme for Information, Dissemination and Training, please contact: **Dr Everardo Sampaio, Coordinator, SIDT, DEN-UFPE, Av. Prof. Luis Freire 1000, Recife, PE, Brasil 50740-540. Email: esampaio@npd.ufpe.br.**

MANAGING BIODIVERSITY DATA WITHIN PNE

Three biological data management systems are currently used by PNE: *Alice*, *BRAHMS* and *DELTA*. These systems have different, complementary functions and are compatible in the sense that it is possible to pass data between them. *Alice* is used for collating, exchanging and disseminating all of PNE's information about species - their nomenclature, distribution, ecology, morphology, use and management. *Alice*, therefore, is at the heart of PNE's information services. *BRAHMS* is used to manage herbaria and specimen collection records while morphological descriptions stored in Alice databases are exported into *DELTA* format to generate identification guides and keys.

Over 20 databases have been built to date containing a wide variety of information in areas such as biodiversity, economic botany, medicinal properties and ecology. Each project leader has ownership and control of the data collated. Each receives technical support and advice on designing their database(s) and selecting software and what data standards are available. Once data have been published and are in the public domain, they are passed to centrally held and coordinated thematic databases. Projects also choose to share most, if not all, of their other data with the common pool. In return, they gain access to centrally held databases, can periodically validate their own data against those of colleagues and can search the whole, asking questions that cross reference aspects of their own research with data from other disciplines.

ALICE DATABASES

14 species databases have been built using Alice:

- medicinal plants (4) c. 950 spp.
- forage plants (4) c. 3200 spp.
- checklists by vegetation type (2) c. 6500 spp.
- regional checklist (1) c. 7000 spp.
- existing data sources (3) c. 3000 spp.

Subsets of a number of key species databases at the Royal Botanic Gardens, Kew are also in the process of being transferred to Alice databases for use by PNE. These include SEPASAL, a database of arid and semi-arid plant species of economic importance, a wood anatomy database and a checklist of *Brejo* vegetation (a Northeast Brazilian forest type).

DATA STANDARDS

PNE established a number of terminological standards to resolve potential incompatibilities in key areas identified as being of common interest. Standards have been designed, for example, for:

- the vegetation types of NE Brazil;
- habit descriptions;
- use properties for forage plants;
- use properties for medicinal plants.

These standards are developed by committees and are not imposed exclusively. Some descriptors are considered essential, some optional and each project can define its own descriptors to capture information of local interest. Some standards have been in use for three years and experience already indicates where changes would be beneficial. Two parallel data standards are planned for medicinal plants, for example, one to be used in ethnobotanical studies and the other in chemical/pharmacological studies. We welcome hearing from Alice users with experience in storing these sorts of data.

DATA EXCHANGE

Since projects build their own databases locally, data are exchanged regularly within PNE. Manually merging large, complex datasets, however, is not for the faint-hearted! Even when datasets have been built using the same program and data standards, inconsistencies still arise (such as for example, if the respective authors use different scientific names for the same plant or have different opinions about its distribution.) *Sam* was a program originally developed by Alice Software in conjunction with the International Legume Database and Information Service Project, another multi-site user of the Alice system. Funding for PNE from the Reuters Foundation enabled further development to overcome some of the limitations of Version 2.0. PNE is now using elements of the new *Sam tool kit*. [Please see Alice Software News 4 & 5. *Ed.*]

For further information please contact Eduardo Dalcin (edalcin@elogica.com.br) or Bob Allkin (ba@di.ufpe.br).

WHY PNE CHOSE ALICE TO MANAGE ITS SPECIES DATABASES

- **Data independence**

Each project leader must control what data are stored in their database(s) and how that data is entered. *Alice* is data independent, giving PNE complete control over its own data and the flexibility to meet present and future needs. Projects can define their own species descriptors and modify these during the life span of the project.

- **Data quality**

To control the validation and integrity of our data, we needed a system designed from the outset to incorporate biological intelligence and biological rules.

- **Data security**

We needed to know that our data could be exported into other programs so that it was safe for the future and giving us the flexibility to use data as we want.

- **Data exchange**

Alice uniquely offered tools through which information arising from two or more databases could be merged into one whilst detecting and resolving data conflicts.

- **Powerful query and reporting tools**

Our information services are not yet defined and we needed flexible options for retrieval and publication, using a variety of media from simple printed lists to distributed data sets on disk with query programs that are easy to use. We certainly need to publish datasets on the World-wide Web.

- **Maturity and reliability**

We preferred to adopt a system that was already tested and in use by others. We certainly didn't have the resources to develop anything new, and needed something that would work from day 1 and that would be maintained.

Bob Allkin moves to Brazil

The plant information centre in the Federal University of Pernambuco will be run by Eduardo Dalcin with a staff of four people. Funding has been assured for the next five years. For the first four years, Eduardo and his team will be joined by Bob Allkin, of Alice Software. Bob will provide technical support as part of the British Government's bilateral aid programme. Bob views this as an exciting opportunity to join an energetic multidisciplinary team exploring the processes by which biodiversity information is gathered, translated and delivered to target audiences. Bob said "My hunch is that while data quality, coordination and flexible database structures will help ensure that all available information can be used to maximum effect, the success of the enterprise, in satisfying genuine user-needs, will depend on the effective two way communication between those generating biodiversity information and those using and working with plants".

A **Windows-based** data capture program for ATF:

Third-party software for the Alice System

Despite its powerful data management capabilities and efficient relational table structure, the suite of programs in the current Alice System lacks a native interface for the Windows operating environment. The advantages of a Windows-based interface can go beyond simply being "current fashion" and include faster data capture (fewer mouse clicks and pull-down menus), ease of use (especially important for novice users) and the possibility of incorporating new datatypes (images and sounds) into a database.

One area in which a Windows-based version of Alice would be immediately useful is in data entry, browsing, and editing. In this context, I am developing a data capture program that allows the creation and editing of data tables in Alice Transfer Format (ATF) that can subsequently be imported into an empty or existing Alice database using *Sam*. This program - which I call *Aline* ("*A*Lice *IN*teractive *E*ditor") - runs stand-alone as a 16-bit Windows application and is compatible with both Windows 3.*x* and Windows '95. *Aline* currently edits data sets in ATF for Alice System 2.0 stored in the xbase/dbase format. While *Aline* allows creation and editing of ATF tables "from scratch", one would, of course, still need a copy of *Sam* to create Alice databases and the Alice System to be able to use the data. The *AliceWeb* program (see Alice Software News 5) generates HTML pages from data sets in ATF and so could be used to "publish" data direct from ATF. *Aline* will generate HTML pages dynamically. All reports produced by *Aline* are already in HTML format, allowing users to publish electronically as data are entered. A possible future enhancement would be for *Aline* to work with ATF for Alice System 2.1.

An experimental version of *Aline* can be made available to any interested Alice user, provided that they wish to collaborate in testing and developing the program further.

Mauro J. Cavalcanti, Departamento de Biologia Geral, Universidade Santa Úrsula, Rua Fernando Ferrari, 75, Botafogo, 22231-040, Rio de Janeiro, RJ, BRAZIL
E-mail: maurobio@uerj.br

Alice and the Windows environment: editor's comments

We applaud Mauro's efforts in developing a Windows-based ATF editor. Its great to see third-party developments (such as *Aline* and *AliceWeb*) developing around the Alice System. Increased use of ATF, with or without Alice itself, can further benefit users and non-users alike.

Release of Alice Systems 2.1 and 3.0 marked the arrival of a user-interface for the Alice System which is consistent for all programs and which makes the best use of ring style menus, on screen forms and pick and point lists. Users accustomed to System 2.0 will be delighted with the ease and flexibility with which data can be browsed and edited in the new Systems.

As Mauro suggests in his article, the benefits which a graphical user interface (GUI) such as Windows programs now normally provide, can be very significant and certainly more than merely cosmetic. Achievement of a consistent and intuitive interface across a sophisticated system such as Alice, however, requires careful planning and design. The Alice team is currently engaged in the design of such an interface which will be used for all our programs, which will benefit from modern object-oriented technologies and which will be suitable for subsequent generations of the Alice System (to manage images, diagrams, bird-song etc). Once this new interface is available we will release new versions of ALL existing programs with a GUI capable of use in Windows95, WindowsNT and other operating systems.

A separate article (page xxx) describes a new program which will manage images within Alice databases. This will be the first program released that uses the new Alice GUI.

Alice Transfer Format (ATF)

ATF was originally viewed as a medium for exchanging data between Alice databases or between an Alice database and other programs or databases. *Brahms* and *Delta,* for example, both generate ATF datasets which can be imported into Alice databases. Publication of the ATF standard generated unanticipated, and growing interest in its use both as a format for third party software (such as *AliceWeb* and *Aline*) and as an exchange vehicle for programs and projects other than Alice users.

ATF is evolving in parallel with the Alice System itself. As new features and data classes are added to Alice so they are included in ATF. A new version of ATF is supported by Alice System 2.1, another will be released for use with System 3.0. With help from Eduardo Dalcin and others work has already begun on the design of ATF for the next generation of the Alice System.

Alice Software is not planning to release an ATF editor such as *Aline,* and we wish Mauro well with his project.

A database of
Rattan palms

Rattans are climbing palms belonging to the Calamoideae, one of the largest subfamilies of the Palm family. We estimate that there are probably over 600 different species of true rattan belonging to 13 genera (climbing palms belonging to other tribes are true rattans).

Much of the trade in rattans, primarily as a source of cane for furniture, relies on plants taken wild from natural forest, and in some particularly diverse areas there may be as many as 30 or 40 species occurring together in the same patch of forest. In Southeast Asia rattans are thought to be the most important non-timber forest product. As such they are worthy of attention, and over the past two and half decades, there has been a dramatic increase in the number of projects based on rattans both in Asian and Africa.

Identification of rattans is crucially important when their use and development potential is considered because, not surprisingly, different species have different cane properties, ecology and biology that can affect their market value and their potential as a plantation crop it is essential that you know exactly which species you are dealing with. Perhaps because they are spiny and difficult to collect they have been somewhat neglected taxonomically until recently. In fact, the astonishing diversity of form and arrangement of spines provides a wealth of features that help species identification, even in the sterile state. It's possible, for example, to distinguish and identify the 100 or more species of rattan indigenous to Peninsular Malaysia from just their leaf sheaths. Many of the species in the Southeast Asian region occur over wide areas, transgressing political boundaries and there has been a tendency for the same species to be described from different areas under different names. Keeping track of all the alternative names of rattans and cross references between them is not straightforward and Alice offered us a perfect tool.

Our ambition has been to cover the entire palm family, and work began in 1988 to build a database. In the early days we were constrained by the computer hardware available to us and, being among the very first users of Alice System 1.0, we also had fun learning the program and discovering some blips. Support from Alice Software was always available, however, to help us out. The main constraint on development of our database has been manpower. We have had a series of students, mostly shared with other projects, and temporary research staff who have had as part of their work, the collation of names into the Alice Palm database. Today we have a database of about 1,450 palm taxa with more than 3,900 scientific names recorded for them along with their distribution and other information.

Thanks particularly to Alison Hoare, Helen Sanderson and Sasha Barrow, we now have a complete coverage of all rattans, the names have been checked and the database edited and expanded to include more detailed information for these species. For other palm groups, the coverage is more patchy, though for some geographical areas coverage is complete.

At present we believe we have all validly published names of rattans in the database. Taxa are designated as accepted or provisional. Names are designated as preferred or synonyms of preferred names. Several dubious names of uncertain application are included as provisional. There are 678 accepted or provisional rattan taxa in the database, including species and infraspecific taxa. As can well be imagined there

is great scope for taxonomic and nomenclatural confusion here, and the database is proving to be extremely useful in answering queries along the lines of "How many rattans have been recorded for Lao?" (surprisingly few, in fact, though on a recent visit to Lao I saw herbarium material indicating that the country has a very rich rattan flora it just hasn't been documented), "Where does *Calamus caesius* occur naturally?", or "What is the current status of the name *Calamus inermis*?"

But this isn't the limit of the database's usefulness. We have begun to use the database as a management tool to develop a DELTA-format data matrix. Using Alice descriptors, Helen, Sasha and I have defined and listed about 90 morphological features that are of use in the diagnosis of species. These have now been scored for the rattans of Brunei and entered into the database. Using *Alex*, we then exported the species descriptions into DELTA format and, with Don Kirkup's assistance, have then used *INTKEY* as an interactive identification key. The prototype seems to work well, though we quickly found out how difficult it can be to record characters that were badly thought out, and, inevitably, some misscoring took place when we first entered the data. This Brunei dataset is being used to develop an electronic account of the rattans of Brunei to parallel the hard copy manual that has just gone to press. With 80 species, the tiny state of Brunei has a rich rattan flora, with many species shared with neighbouring Sabah and Sarawak, and even Kalimantan. It will not require much extra effort to extend the dataset to include the whole of the island of Borneo.

While this is going on, two other datasets are being developed. Tom Evans, employed on a Darwin Initiative-funded project on the rattans of the Lao People's Democratic Republic which is jointly managed by John Dransfield and Nick Brown in Oxford Forestry Institute, has started to score the rattans of Lao, Cambodia, Vietnam and neighbouring parts of China, Burma and Thailand, aiming to produce an interactive key to the rattans of Lao. Rudi Maturbongs from Universitas Cenderawasih in Manokwari, Irian Jaya, Indonesia, is visiting Kew for a six month period and plans to enter description for the rattans of New Guinea, again using Alice. In both cases the names of the rattans are already in the database and the descriptors already defined.

Alice has proved to be an excellent program for managing the names, distributions and descriptions of palms and we have been able to "design" our database to suit our own ends. One area that we adapted early on was the use of bibliographic citations for rattan distribution records. Rattan distributions, in fact, are difficult to cull from the literature, and, even in the instance of commercially important material, there may be no published record for the occurrence of the species in a particular area. We simply inserted the collector and number of a herbarium specimen in the reference field to act as a data source.

Recently the Alice-Web program (see Alice Software News 5) was used to create sexy Web Pages, complete with images, from a subset of our data. The potential is fantastic and, with a growing interest outside Kew in the information we have in the database, I cannot wait to get the whole up and running on the Web and accessible to the widest possible audience.

John Dransfield
Herbarium, Royal Botanic Gardens, Kew, Richmond, Surrey, TW9 3AB.

New CITES Orchid Checklist publication

The second volume of the CITES Orchid Checklist has been published by the Royal Botanic Gardens, Kew. The series of Checklists, which are generated from Alice databases, provide details of orchid genera identified as priorities in the CITES 'Review of Significant Trade in Species of Plants: Appendix II" list.

This second volume contains details on the genera: *Cymbidium*, *Dendrobium* (some sections), *Disa*, *Dracula* and *Encyclia*. The book consists of three parts:

1. An alphabetical list of all preferred names (1,234) and synonyms (2,173) included in the checklist - a total of 3,407 names.
2. A listing for each genus, in alphabetic order, giving full details on its current synonymy and distribution.
3. A list of taxa recorded for each country in which these Orchids occur.

Roberts, J.A., Allman, L.R., Beale, C.R., Butter, R.W., Crook, K.R., McGough, H.N. (1997). CITES Orchid Checklist Volume 2. pp 312.

Further enquiries or copies from:
Jacqueline Roberts, Conventions and Policy Section,
Royal Botanic Gardens, Kew. Richmond, Surrey TW9 3AE. UK
Tel: +44 (0)181 332 5724, Fax: +44 (0)181 332 5757

Aboriginal biological knowledge: its conservation a

The primary aim of the project is to conserve and promote the diversity of traditional Aboriginal knowledge of the plants and animals in the wet dry tropics of the Northern Territory (NT), Australia. We work, where requested by Aboriginal people, in a scientifically sound and culturally sensitive manner.

The importance of conserving and promoting biological knowledge has been recognised by Aboriginal people for a very long time, as exemplified by the number of related ceremonies, significant or sacred sites and totems. The importance of traditional biological knowledge was more recently recognised internationally by UNESCO's Man and the Biosphere programme. Conserving ethnobiological knowledge within Australia is one of the key objectives in the National Strategy for the Conservation of Australia's Biological Diversity (1995).

THE PAST

An ethnobotany project began in the Northern Territory Herbarium, Darwin in 1991. In response to repeated requests from Aboriginal groups, a botanist and a trainee were given the task of conserving and promoting traditional plant knowledge. While some ethnobotanical work had been undertaken prior to this, the allocation of full time staff to the initiative, signalled a shift in priorities and allowed an enormous increase in the conservation of traditional plant knowledge.

Our primary mechanism for conserving knowledge is through the NT Botanical Bulletin series which details traditional plant knowledge of individual language groups. To date, Bulletins have been produced for 11 Aboriginal language groups. Production of the last two Bulletins have been primarily funded by the communities with whom they were produced.

Further dissemination and promotion of traditional plant knowledge has been via posters, identikits, plant use walks at the Darwin Botanic Gardens and larger thematic books. A series of seven posters relating to different plant uses, in conjunction with a series of four Identikits, relating to "Top End", "Desert Bush Tucker", "Bush Medicine" and "Jawoyn" plant knowledge, have been particularly popular. "Top End" is the local term for the wet-dry tropics area of the Northern Territory. The "Jawoyn" people are the traditional owners of an area that includes one of Australia's best known National Parks: Nitmiluk. "Bush Tucker" refers to plant and animal foods used by Aboriginal people and collected from the wild. Aboriginal plant use walks at the Darwin Botanic Gardens have been very popular with the general public, special interest groups and schools. More thematic, self-guided walks are being introduced as the Botanic Gardens new strategy is implemented. A central theme for the Botanic Gardens is the use of plants by man. Publication in 1993 of a large format book entitled Traditional Aboriginal Medicines, symbolised the Commission's commitment to cross cultural research and conservation.

THE PRESENT

In response to requests from Aboriginal people we began recording animal names and uses in 1995. Currently staff of the Ethnobiology Project are working with elders from sixteen Aboriginal language groups in the "Top End" region, recording the traditional knowledge associated with the native plants and animals. It is hoped that two or three of these accounts will be completed in 1998.

Requests for assistance continue to be received from Aboriginal language groups. These requests are often prompted by a realisation that most of the more knowledgeable people within any community are now elders. The passing of such elders will represent a quantum loss of knowledge from the community. The latter recognise that the current generation of senior knowledge custodians are probably the last bearers of such an extensive and detailed understanding of the biota which was gained through living in and off the country for extensive periods as youngsters, under the tuition and guidance of their own parents and elders.

THE FUTURE

The impending significant loss of traditional Aboriginal biological knowledge over the next ten years poses our project with a severe problem. Many more language groups already seek our assistance than current staffing levels can handle. Requests have come, however, only from those language groups that are already aware of the Commission's project. Were we to actively solicit interest then it is likely that most Aboriginal language groups in the Northern Territory, between 30 and 50 depending upon definition of a language or dialect, would wish to participate.

We hope that further funding will allow the Ethnobiology Project to continue and to expand its role in assisting Australia's neighbours in recording plant knowledge. Co-operative research and training have been undertaken in Indonesia, the Solomon Islands and New Guinea. Recent funding from the Australian Centre for International Agricultural Research will be used to undertake field work and training in eastern Indonesia with staff from the Indonesian Botanic Gardens.

Glenn Wightman (for address please see below)

nd promotion

THE ETHNOBIOLOGY PROJECT'S DATABASE

Our project has used Alice to store and query data since 1989. We currently use System 2.0 and plan to move shortly to System 2.1. Our database contains information about over 1100 native plants recorded with their various Aboriginal names and the uses to which they are put in the Northern Territory. Data from over 40 Aboriginal language groups has been included so far. Information has been collected from a wide range of sources and aims to draw together and coordinate knowledge into one resource. Descriptors of use, ecology and morphology have been based on international standards where appropriate and possible. A detailed checklist of plant names and uses in the Northern Territory will be generated from the database for publication as text (using *Awrite*) or on the WWW (*AliceWeb*).

The Alice system has provided a convenient way to store and query our plant names and recorded uses. Data storage has been reliable and backup support from Alice Software has been good. The ease and flexibility with which new descriptors can be set up and the System's database maintenance facilities have both been important to our work.

Nick Smith (for address please see below).

Full address for authors:
Parks and Wildlife Commission of the Northern Territory, Australia. PO Box 496, Palmerston NT 0831, Australia;
phone + 618 8999 4513,
fax + 618 8999 4527,
e-mail glenn.wightman@nt.gov.au

Database Statistics (as at April 1998)

Total no. of Species	1,144		
Aboriginal names	8,973	Project descriptors	12
		Recorded observations for descriptors	1,9270
Scientific names	1,144	Geographic distribution records	10,111
Scientific synonyms	150	Literature pointers for taxa	367
Misapplied names	61	Text notes for taxa	6,030
Orthographic variants	23	Uses recorded for taxa	1,814

Conservation Commission of the Northern Territories Ethnobiological database

Descriptors	No.States
Aboriginal community where recorded in use	93
Economic importance by ethnic group	3
Language name	50
Use: Food type	14
Use: Implement or weapon type	15
Use: Materials type	18
Use: Medicinal ailment	22
Use: Medicinal treatment type	10
Conservation status	4
Endemic status	2
Introduction status	2
Lifeform	25

Example

Descriptor:	**Use: Implement or weapon type**
Definition:	Unordered / Variable
	States
1	Axe
2	Boomerang
3	Clap-stick
4	Coolimon And Carriers
5	Didgeridu
6	Digging-stick
7	Fighting-stick
8	Fire management
9	Message-stick
10	Pipe (smoking)
11	Shield
12	Spear
13	Water craft
14	Woomera
15	Other

Indexing the world's known species

The Species 2000 Programme has the ambitious aim of creating a uniform and validated index to the world's known species for use as a practical tool in inventorying and monitoring biodiversity worldwide. This index will be formed by linking a number of individual taxonomic databases to form a virtual "Dynamic Checklist" on the World Wide Web and a standard "Annual Checklist" produced on an annual basis, available on CD ROM as well as the Internet.

The ILDIS World Database of Legumes, which is based on Alice software (see Alice software news 5) is one of those being used as a prototype in the initial development of the project. It is hoped that other Alice databases will also play a role in the future and interested parties are invited to contact the Species 2000 Secretariat (below).

THE SPECIES 2000 PROJECT

The names of organisms are the key to biodiversity communications and as such, provide access to the accumulated knowledge of all life on Earth. It is remarkable that to date no comprehensive indexing system exists for the 1.75 million plants, animals, fungi and micro-organisms named by science.

The *Species 2000* index will be used to provide:

- electronic baseline species lists for use in inventorying projects worldwide;
- the index for an Internet gateway to species databases worldwide;
- a reference system for comparison between inventories;
- a comprehensive worldwide catalogue for checking the status, classification and naming of species.

The central concept of the *Species 2000* programme is the array of taxonomic databases, each covering a group of organisms at species level, referred to as Global Species Databases (GSD's) in the diagram. This array is organised by one or more taxonomic hierarchies connecting the groups of organisms (see diagram).

The taxonomic database organisations starting the Programme already provide or are developing such indexes for: Viruses, Bacteria, Corals, Molluscs, Crustacea, Flies, Ichneumon Wasps, Geometrid Moths, Weevils, Fishes, Birds, Mammals, Fungi, Mosses, Fossil Plants, and Angiosperms (including Cacti, Palms, Legumes and Umbellifers). Organisations with databases covering a further 40 major groups are joining the Programme and it is projected that existing database projects may provide for approximately 40% of known species. One hundred or more global species databases, each initially covering 10,000-25,000 species, will be needed for all species to be included. *Species 2000* proposes to stimulate completion of the array of taxonomic databases. It will seek resources both for the completion of the existing databases, and to help establish new databases to cover the gaps, thought to account in total for about 60% of species.

A further important development is the provision of onward species links as part of the Internet gateway. Once a species has been located, onward links are provided to rich data sources for that species in a variety of conservation, taxonomic, germplasm or ecological databases in different countries.

A prototype (see illustration) tested on the Internet demonstrates the feasibility of addressing an array of data-

Species 2000 prototype user interface for accessing an array of databases on the Web

bases in this way. Checklist elements of ILDIS LegumeWeb, Prokaryobase, FishBase and the Plant Fossil Record are accessed via a single interface. Development of the working system and technical work on interrelating existing databases is under way.

Species 2000 is becoming a large international co-operative programme. It was established by the international scientific unions IUBS, CODATA and IUMS, subsequently endorsed by UNEP in its Biodiversity Work Programme 1996-1997, and associated with the Clearing House Mechanism of the UN Convention on Biological Diversity. It is now registered as a not-for-profit international organisation managed by a project team responsible to the participating databases.

Three partner programmes participating in *Species 2000* are: ETI's UNESCO/IOC *Register of Marine Organisms*; the North American *Integrated Taxonomic Information System* (ITIS) and the *Global Plant Checklist* of the International Organization for Plant Information *(IOPI - see separate article).*

ALICE DATABASES AND SPECIES 2000

There are two modes in which Alice species databases are expected to participate in *Species 2000*. The role of ILDIS LegumeWeb as one of the GSD's in the primary array may be followed by other Alice-based taxon databases: for example the Royal Botanic Gardens Kew has Alice-based databases for the Palms and the Orchids. Secondly, we can envisage other Alice-based taxonomic databases such as those on crop cultivars, regional species inventories and ethnobotany being linked to Species 2000 in the secondary array by onward species links.

Frank A. Bisby and Sue Tong

Webpage:
in Europe:
http://www.sp2000.org
in Asia:
http://www.japan.sp2000.org
in North America:
http://www.atcc.org/sp2000/

Secretariat:
Species 2000 Secretariat, Centre for Plant Diversity and Systematics, School of Plant Sciences, University of Reading, Reading RG6 2AS, UK.
Fax: +44 118 975 3676.
Tel: +44 118 931 6466.
Email: sp2000@sp2000.org

Sponsors:
IUBS: International Union of Biological Sciences
CODATA: ICSU Committee on Data in Science and Technology
IUMS: International Union of Microbiological Societies

Alternative Editions of the Alice System

Alice Systems 2.1 and 3.0 are supplied as one of a number of alternative "editions". Each edition contains a different subset of the available programs, according to the needs of different types of user.

- The **Database Author's edition** is for those who build their own databases.
- The **End User's edition** provides 'read-only' access to Alice databases. Database authors can thus publish their databases (or subsets of them) electronically by distributing them with copies of the End user's edition, enabling colleagues to have controlled access to their data.
- The **Network Manager's edition** is for use by the network administrator in an institute where several users work with more than one Alice database.
- The **Database Administrator's edition** is for use by the data administrator in an institute, or collaborative projects where several users are working with Alice databases and aim to share their data, their procedures and common data standards.

Programs available with different Editions of System 2.1 or 3.0

EDITION	Ace	Aview	Aquery	Awrite	Alice-Web	Alex	Sam Tool Kit	Amie	Db-update	Db-admin	Db-unlock
Database Author's	•		•	•		•		•	•		•
End User's		•	•								•
Network Manager's	•		•	•		•		•	•	•	
Database Administrator's	•		•	•	•(*)	•	•(*)	•	•	•	

(*) This program is, as yet, unavailable for System 3.0.

The Millennium bug: "Y2K Compliance"

For those that haven't come across the term, "Y2K" is an acronym for "Year 2000" and is used to refer to a problem that is increasingly pre-occupying IT managers around the world. Their concern is whether software, computer systems and any other device containing micro-processors will manage dates properly in and after the year 2000. We can't discuss all the technical details here, but it is important for users of Alice to know whether Alice is Y2K compliant or not. Some organisations buying software are now insisting that software purchased is Y2K compliant and may also seek assurance that programs previously purchased are, or will be, Y2K compliant.

WHAT ABOUT ALICE?

Alice Systems 2.0, 2.1 and 3.0 are not Y2K compliant according to the very strictest of definitions in use but for most purposes, and using most definitions of compliance, all Systems are already Y2K compliant. Let us explain.

THE GOOD NEWS

Alice Systems perform no calculations that use dates. The only context in which dates are stored in Alice databases is in the citations manager, where they are stored as a four digit field, which should see the software through to the year 9999! The underlying development language in which Alice is written is Y2K compliant and there is no reason to suspect that any Alice program will fail to execute or cause any kind of system failure in or after the year 2000.

THE BAD NEWS

Alice does display the current date, however, particularly in reports and on certain screens, using only two digits not four. By the strictest of definitions therefore Alice could be said to be not Y2K compliant. The current date could easily be displayed as four characters, however, and the necessary modifications to the software are being made.

Y2K COMPLIANT VERSIONS

Before the end of 1998, updates to System 2.1 and System 3.0 will be released that, along with various other, more significant advances (please see separate articles), will contain the necessary modifications to displaying dates. More significantly, these new releases will also have been tested using industry standard tests to ensure that BSI (British Standard Institute) standards are met. These new upgrades will be formally certified "Y2K compliant". No such upgrade is planned for System 2.0.

As always, new releases will be made free to all users with current maintenance agreements. Other users will be asked to pay an upgrade fee.

A FINAL WARNING

The dates used by Alice are read from the date as understood by your computer. If your computer is not Y2K compliant, then you may still have problems.

NB. A full definition of Year 2000 Conformity can be obtained from the British Standards Institute, 389 Chiswick High Road, London W4 4AL.

HOW TO REACH ALICE SOFTWARE

We urge all of our readers to correspond with Alice Software via e-mail. We cannot guarantee a rapid response to letters or faxes. Our email address is ***alice@dial.pipex.com.***

The latest information about the Alice Systems can be found on our World-wide web site which has been recently updated. ***http://dspace.dial.pipex.com/alice***

If you can only write to us by ordinary mail then please write to the following address and clearly label the envelope for attention of Alice Software.

Dr. Bob Allkin (for the attention of Alice Software)
CEB, Herbarium
Royal Botanic Gardens Kew
Richmond, Surrey TW9 3AB, UK
Fax: +44 - 181 - 332 - 5740

NEXT ISSUE

What don't you like about Alice? What do you long for Alice to do? Why is your project a success? Tell us what **you** think!

The next issue of ASN will, amongst other things, contain a description of the ways in which you can now search for and edit names, explanations of the new categories of name and taxon and other datatypes such as introductory status. We have been promised contributions from a large conservation checklist project and from an innovative project about the Loranthaceae (the mistletoe family) which links specimen data and mapping facilities to an Alice database.

Any articles for inclusion in the next issue should reach us (on a disk or by email) *no later than the 1st of Dec. 1998 please.* Let us know what you - the reader whoever you are - would like to see in a future issue.

Alice software news

EDITORIAL AND DISTRIBUTION POLICY

Alice Software News (ASN) is edited by Bob Allkin and Peter Winfield. It is distributed free of charge to users of the Alice System and others that wish to be informed about developments of the System.

Anyone may send articles or letters to be considered for publication in ASN. We reserve the right to edit and reduce contributions. Any changes to text will be agreed with the author before publication.