



BioLink 2.0

A preliminary evaluation

Arthur D. Chapman
September 2003

1. Aim:

To evaluate BioLink Version 2.0 and assess its value for use as a SpeciesLink database.

2. Caveat

Much of the information cited here is taken from the "About BioLink" documents accompanying the software (CSIRO 2003).

3. Background:

BioLink (Shattuck and Fitzsimmons 2000) is "integrated software for the collection, maintenance, analysis, application and dissemination of taxonomic, biodiversity and environmental information" (CSIRO 2001). It is primarily designed to assist researchers and curators working with taxon and specimen-based information such as taxonomists, ecologists, collection managers and biogeographers. To quote from the Web page (CSIRO 2001).

"BioLink manages taxon-based information such as nomenclature, distribution, classification, ecology, morphology, illustrations, multimedia and literature. Specimen-based information (the "Who, What, When and Where" information) includes collection sites, collectors and collection dates, museum storage locations, loans and accession and catalogue numbers. For those involved in surveys, BioLink records information on ecological characteristics, traps and bulk samples."

Major components include:

- A taxon management module
- A module for managing sites, site-visits and collection material
- A module for managing specimen and morphological descriptions
- An on-line electronic gazetteer (eGaz) - also available as stand-alone product.
- A Mapping module for making simple distribution maps
- An Import Wizard
- A loans management module
- Multimedia capture and storage module
- A reference manager
- A report manager
- A query tool
- On-line and text-based help.

Each of the components are described in more detail further on in this report.

4. Technical Details:

BioLink is a Microsoft Windows based system running on Windows 98/NT/2000 and XP. It can be run as a stand-alone system or across a Local Area Network (LAN).

BioLink is developed as a three-tiered databasing application using Microsoft Visual Basic Ver. 6 (primarily) and Microsoft Visual C++ (secondarily) for the user-interface/client layer and middle tier and the Microsoft Data Engine (MSDE) version of SQL Server 2000 as the database layer.

The BioLink installer is written in Inprise (Borland) C++ Builder. A number of third-party custom controls are used, including ESRI's MapObjectsLT, Bennet-Tec's TList and AllText, Apex's DBGridPro and Pegasus' Imagn' as well as Microsoft's Internet Explorer.

The client layer interacts with the middle tier using COM/DCOM while the middle tier and database communicate using the Microsoft's ADO (with options to use other protocols if required).

Information in SQLServer is maintained and accessed using stored procedures and triggers rather than direct access to the underlying tables.

BioLink Ver. 2.0 contains approximately 400,000 lines of Visual Basic code, 15,000 lines of C++ code and the database contains approximately 730 fields in 77 tables, 650 stored procedures and 95 triggers. (CSIRO 2003)

5. Licensing:

The software has an open and generous licensing agreement, allowing the user to use the product on any number of computers and/or networks simultaneously; make any number of backup copies, and freely transfer the product as long as no fee is charged and the copyright statement is also transferred.

6. Evaluation:

6.1 Installation:

Installation is simple and easy, using an Installation Wizard. As part of the installation process, Microsoft's Database Engine (based on SQL Server 2000) is also installed.

6.2 Demonstration Database:

The program comes with a demonstration database with included collections, along with an "About Biolink" help file that gives an Introduction and explains each of the modules.

6.3 Modules:

6.3.1 Taxon-based Information

The Taxon-based information provides information on whole taxa rather than individual specimens. It includes such information as:

- Valid (zoology)/Accepted (botany) Name
- Available Name (zoology)/Basionym (botany)

- Literature (zoology)/Valid (botany) Name
- Classification
- Nomenclature/Type Data
- Common Names
- Political Distribution
- Literature References
- Images/Sounds/Multimedia/Digital Information
- Associates
- Keywords
- Storage Location
- User-defined Traits and Notes

The Taxon module has a number of “Explorers” that display the names of all the taxa held in the BioLink Database, together with their classification; Distribution Regions; and Storage Locations. There are three tabs in each of the Explorers (fig. 1).

- Contents – contains a complete expandable hierarchical listing of all taxa in the database (fig. 1)
- Find – used to locate names in the database using a simple search
- Favorites – for organising commonly used names or sets of names.

The first Explorer is for names of taxa held in the database along with their hierarchical classification (fig.1).

Internally, BioLink stores the names of taxa separately from author and year of publication. When one enters a name, however, it is entered as a string, and BioLink then attempts to parse them into different fields. It is important to check that the name has gone in correctly. By right-clicking on a name and selecting “Edit taxon” a dialog box pops up and allows one to edit the record

Further details may be added to the Taxon information by Right-clicking on the name and selecting “Edit Taxon Details”

This allows one to add information such as:

- Common Names
- References
- Distribution
- Multimedia (images, video etc.)
- Associates
- Traits
- Keywords
- Notes
- Location

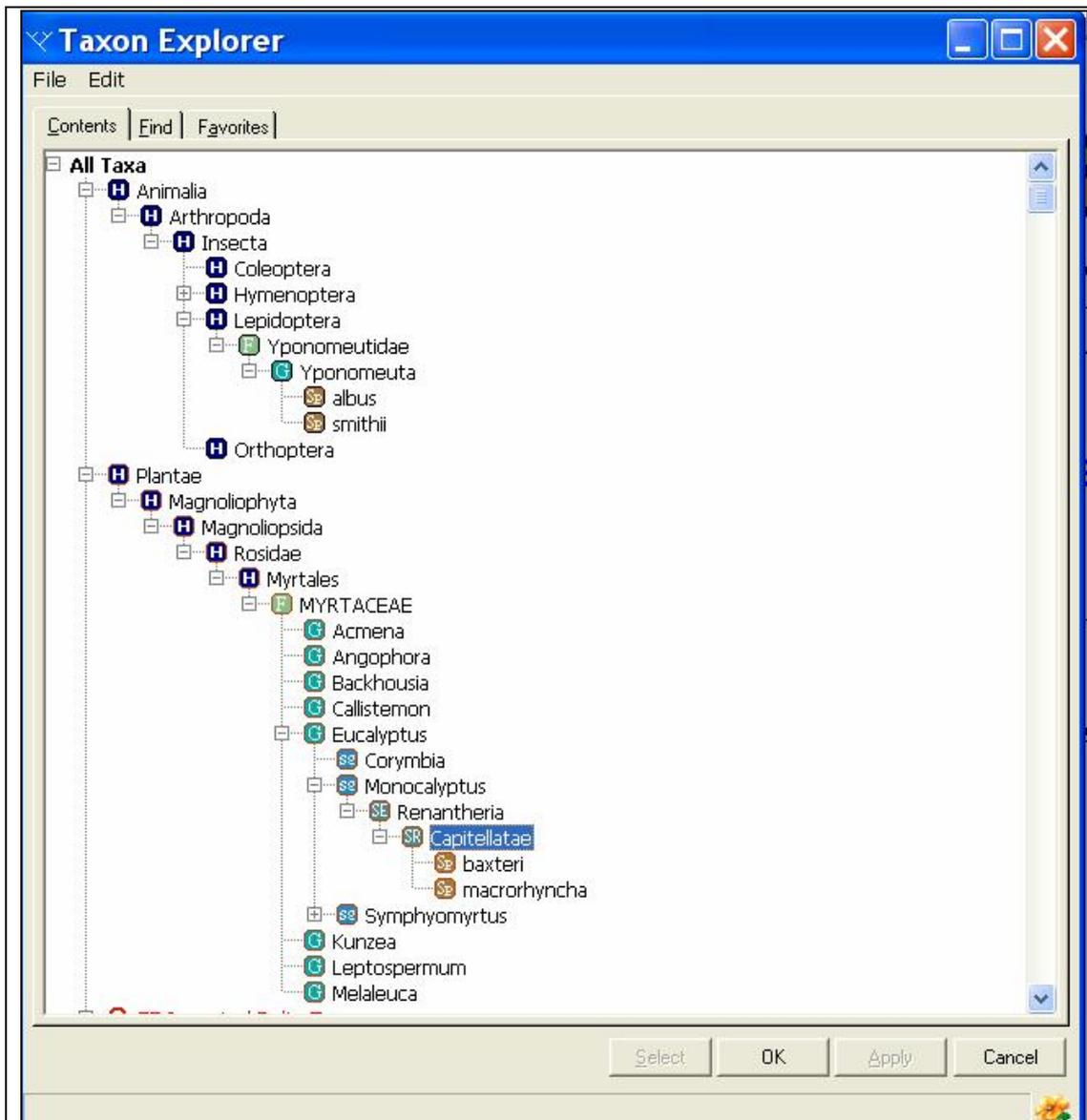


Fig. 1. The Taxon Explorer from BioLink, showing an expanded hierarchy for one species of insect and two species of *Eucalyptus*.

The second Explorer is for Geographic Regions, again laid out in an hierarchical fashion. As for taxa, it is easy to add new regions, for example, States. The hierarchy can be taken down to as many levels as one may wish.

The third Explorer is for storage locations of the collections. Again, this is in the form of a hierarchy, and can be taken down to the level the user desires – through collection, room, cabinet, draw or whatever. Again – it is easy to add new Storage localations.

Evaluation: Using the Contents, I found it easy to add a new name, geographic region or storage location into the hierarchy at any level. Once having read the instructions, it was quite intuitive.

To change the position of a taxon in the hierarchy, it is just a matter of dragging the taxon from one position to another.

Evaluation: Using the Contents, I found it easy to modify the position of a taxon in the hierarchy, it was quite intuitive.

6.3.2 Specimen-based Information

The specimen-based information is for managing collections of observations of specimens. It is commonly associated with museum of herbarium-based material. It includes such information as:

- *Collection Sites or Localities*
 - Political Region
 - Locality Description
 - Geographic Position (Latitude/Longitude or UTM/Map Grids)
 - Elevation or Depth
 - Geological Details
 - User-defined Traits and Notes
- *Collecting Events*
 - Collectors
 - Collector's Field Number
 - Dates and Times (Single or Period)
 - User-defined Traits and Notes
- *Specimens/Observations*
 - Accession, Registration and Collector's Numbers
 - Collection Method
 - Macro- and Microhabitats
 - Abundance
 - Institution
 - Identification (Current and Previous)
 - Number of Specimens, Gender and Life Stage
 - Specimen Subparts and Multiple Specimens
 - Original Label Text
 - Storage Location and Storage Method
 - Condition and Curation Status
 - Associates
 - Curation Events
 - Specimen Label Details
 - User-defined Traits and Notes
- *Traps and Bulk Samples*

The specimen-based information can also be displayed and edited through a "Site" Explorer. The Site Explorer displays sites, site visits and material held in a BioLink database. It is organised hierarchically by region, countries, however this can be user-defined. This data is arranged via a collection event (fig. 2).

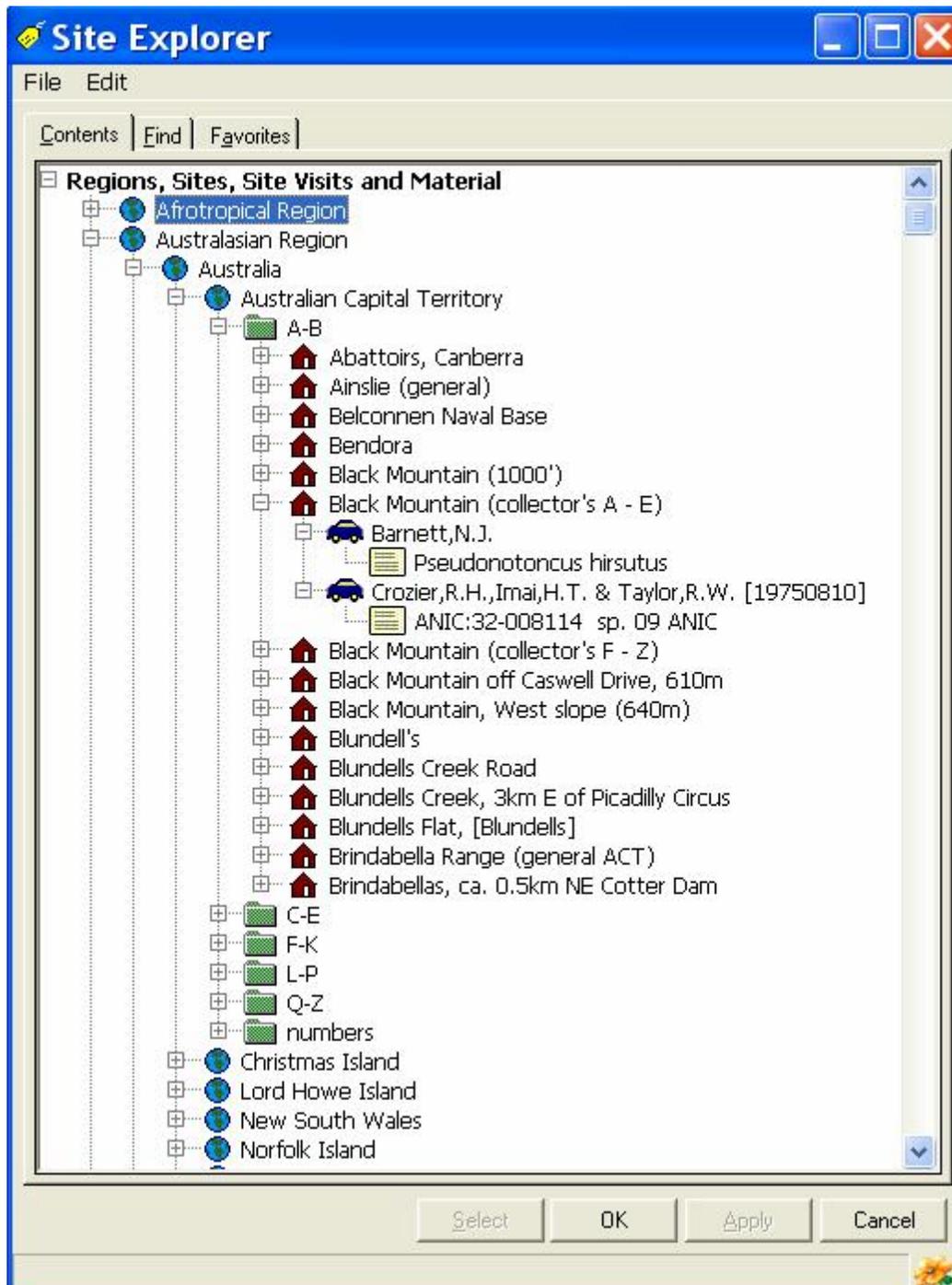


Fig. 2 The Site Explorer showing an expanded hierarchy and two collection events.

By right clicking on a collection event, one has the option of doing a simple edit of the displayed information, or bringing up the “Rapid Data Entry Form” (fig.3).

Fig. 3. Rapid Data Entry Form from BioLink.

By clicking on any of the link buttons (...), one links to one of the other Explorers mentioned above (Taxon, Geographic Region, etc.) or pick lists (List of Collectors, etc.), or Zoom buttons (🔍) which links to one of the editing forms for that field mentioned above.

By clicking on the Link Button (...) beside the Longitude field one brings up the eGaz electronic Gazetteer (see discussion below).

Evaluation: I found this form quite intuitive and to easy use. Once the Taxonomic Hierarchy, Geographic Regions and other Pick lists etc. are loaded, it was quite quick and simple to use. It is oriented toward animal, and especially insect collections, but this does not mean that it is not also valuable for other groups of taxa.

6.3.3 Morphology-based Information

Morphology-based information is used to describe specimens or taxa in the form of a data matrix, and also allows for the development of dichotomous keys. This is a new module first introduced into this version of the software. Descriptions are coded by combining character information (states or values) and comments for each entity, one character at a time following the Delta format (Dallwitz 1998, Dallwitz *et al.* 1993). The Morphology Explorer allows the user to visualise and edit the information

Figure 4 shows a small part of a grass data matrix developed from Watson and Dallwitz (1992).

Morphology Project Matrix - Grasses [ID:2]

Matrix Edit Tools Options

Entities	longevity of plants	reeds	habit	the flowering culms ...
<i>Acamptoclados</i> <Nash>	perennial with remain...		caespitose	
<i>Achlaena</i> <Griseb.>	perennial with remain...		caespitose	
<i>Achnatherum</i> <Beauv.>	annual or biennial, wi...		caespitose	
<i>Aciachne</i> <Benth.>	perennial with remain...		caespitose	
<i>Acidosasa</i> <Chu and Chao>	perennial with remain...			leafy
<i>Acostia</i> <Swallen>	perennial with remain...		caespitose	
<i>Acrachne</i> <Wright & Arn. ex Chiov.>	annual or biennial, wi...		caespitose	
<i>Acritochaete</i> <Pilger>	annual or biennial, wi...		decumbent including ...	
<i>Acroceras</i> <Stapf>	annual or biennial, wi...		conspicuously rhizom...	
<i>Actinocladum</i> <McClure ex Soderstr...>	perennial with remain...		caespitose	leafy
<i>Aegilops</i> <L.>	annual or biennial, wi...		conspicuously rhizom...	
<i>Aegopogon</i> <Humb. & Bonpl. ex Will...>	annual or biennial, wi...		conspicuously rhizom...	
<i>Aeluropus</i> <Trin.>	perennial with remain...		conspicuously rhizom...	

Views Traits M.M.

Views Description [Interactive Identification, *Acamptoclados* <Nash>, generic name author] - Text

Interactive Iden...

Multimedia

Name	File Info

Add New Delete

Fig. 4. A part of a grass character matrix from BioLink. Data from Watson and Dallwitz (1992).

From the character matrix, descriptions of the entity can be built.

Dichotomous Key Builder - Australian Ant Subfamilies [3]

File Edit

Interactive | Text

Start

- Mesosoma attached to the gaster by a single,
 - Gaster with a slight but distinct impression
 - Upper surface of the tip of the gaster
 - Cerapachyinae
 - Upper surface of the tip of the gaster
 - Ponerinae
 - Gaster uniform in outline, without an impr...
 - Mandibles elongate and straight, with
 - Ponerinae
 - Mandibles triangular, with distinct tee
 - Sting present and usually visible at
 - Sting absent (Fig. 30aa), petiole e
 - Tip of the gaster with a circul
 - Formicinae
 - Tip of the gaster slit-like and r
 - Dolichoderinae
- Mesosoma attached to the gaster by two segr
 - Mandibles very long, more or less straight,
 - Ponerinae
 - mandibles generally triangular (Fig. 36aa), l
 - Eyes almost always present, conspicu
 - First segment of the mesosoma (p
 - Pseudomyrmecinae
 - First segment o the mesosoma (p
 - Myrmicinae
 - Eyes absent (Fig. 39aa) or at most wit
 - Antennae with ten segments (incl
 - Aenictinae
 - Antennae with 12 segments (incl
 - Leptanillinae

Tip of the gaster with a circular or semicircular opening (acidopore) which is often fringed with short hard (Figs 31a, 32a) (edit text...)

Tip of the gaster slit-like and never with a fringe of short hairs (Figs 33aa, 34aa) (edit text...)

Alter Image...

Alter Image...

Taxa: Formicinae;

Taxa: Dolichoderinae;

OK Apply Cancel

Completed loading elements.

Fig. 5. An example of an illustrated dichotomous key from BioLink's Morphology Explorer.

The Morphology Module allows the user to develop new character matrices, and dichotomous keys (fig. 5) using the Dichotomous Key builder (Dallwitz, *et al.* 2000), explore the character states, etc. It also allows the import and export of files in both Delta (Dallwitz 1980, Dallwitz *et al.* 1993) and Lucid (Thiele *et al.* 1997) formats.

The Morphology module allows for the inclusion of images and other multimedia information.

Evaluation: Without data to fully explore this module it is difficult to fully evaluate. It is a powerful tool however, especially with its links to data in both Delta and Lucid formats. This module takes the program away from just being a specimen management system and gives it a much more expanded role. Needs further evaluation.

6.3.4 Collection Management Information

This module is for use by museum and herbarium collection managers for managing and curating their collections. It includes tools for managing information such as:

- Accession, Registration and Collector's Numbers
- Taxon and Specimen Storage Locations
- Specimen Storage Method
- Specimen Condition and Curation Status
- Specimen Curation Events
- Lot and Bulk Sample Details
- Identification History
- Specimen Label Details
- Taxon and Specimen User-defined Traits and Notes
- Loans
- Loan and Permit Numbers
- Borrower, Receiver and Authoriser Details
 - Name and Title
 - Job Title
 - Institution
 - Street and Postal Address
 - Work/Home/Fax Phone Numbers
 - Email Address
- Transfer Method
- Date Initiated and Date Due
- Partial Returns
- Loan Conditions
- Description of Material (Specimen- and Taxon-based)
- Automatic Loan Forms
- Correspondence Record
- Reminders and Action Dates
- User-defined Traits and Notes
- Label Manager

This was the original purpose of BioLink when first written.

The information is handled through the Site Explorer and collections module (above) and the Loans module.

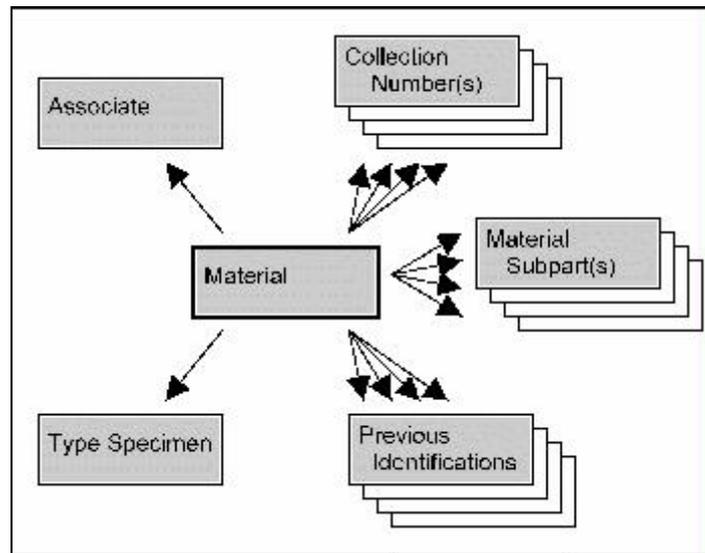


Fig. 6. An example of the links in the Material module of BioLink.

The loans module allows users to keep track of material and correspondence associated with the loan and includes reminders which appear in the Reminders window on user-set dates.

Evaluation: I found it a little difficult to find the place one would track some of this information. For example, it is not intuitive where one would be able to find the identification history. The loans module, however, is intuitive and looks to be easy to use. The collection information is oriented towards entomological collections, but should be suitable to other collection types – especially faunal collections.

6.3.5 Import/Export

An Import/Export Wizard allows data to be imported or exported to/from BioLink in Delimited text, Excel Spreadsheet, DBF or Microsoft Access formats. There is also an XML Data Transfer facility as well as the ability to Import and Export character data in DELTA and LUCID Formats as mentioned above.

Evaluation: I have not attempted to import or export data from the program.

6.3.6 EGaz and Geocoding

The BioLink software comes with the Gazetteer program EGaz included. eGaz is also available separately by downloading the stand-alone software from CSIRO at <http://www.biolink.csiro.au/egaz.html>.

eGaz (Shattuck 1997) is a program developed at the CSIRO's Australian National Insect Collection to assist museums and herbaria to identify and add geocodes to their specimen records.

eGaz eliminates the need for paper based maps and rulers to determine the latitude and longitude for cities, towns, mountains, lakes and other named places. eGaz can also calculate latitude and longitude for sites a known distance and direction from a

named place. The program allows for the easy inclusion of Gazetteers from any region, and Gazetteers for much of the world are available for download from the CSIRO site (<http://www.biolink.csiro.au/gazfiles.html>).

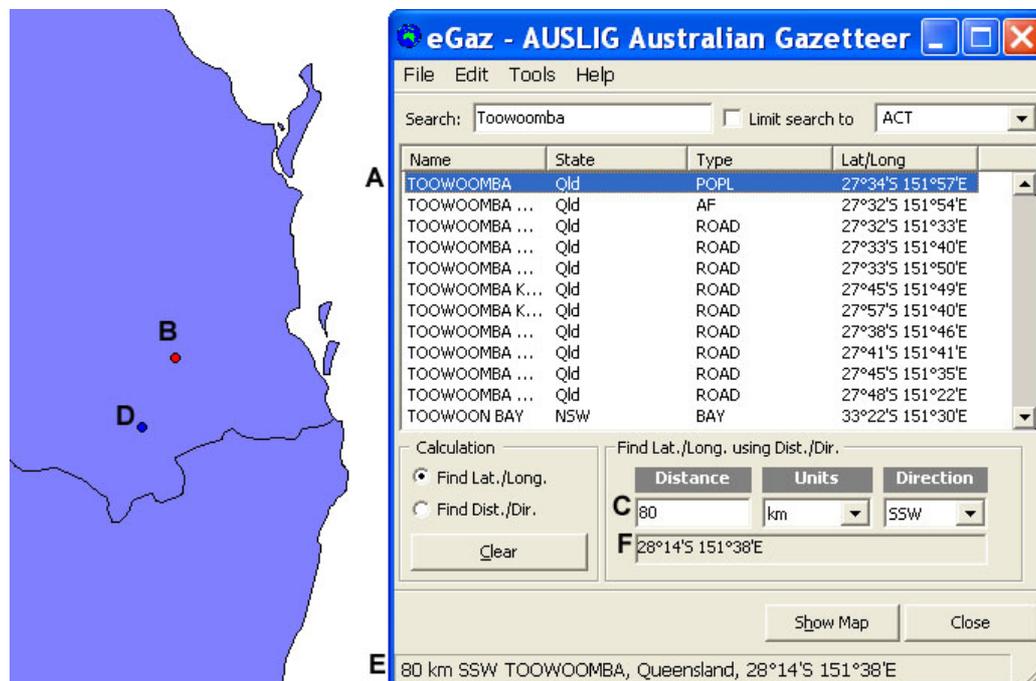


Fig. 7. Sample output from eGaz (Shattuck 1997), showing the determination of latitude and longitude for a position 80 km SSW of Toowoomba, Queensland, Australia. **A.** Information on Toowoomba from Gazetteer. **B.** Mapped location of Toowoomba. **C.** Input showing 80 km SSW of highlighted location. **D.** Mapped location 80 km SSW of Toowoomba. **E.** Details on location. **F.** Latitude and Longitude of new location.

eGaz is a Microsoft Windows based product that provides two windows, a Gazetteer window and a Map window (fig.7). It allows the user with a location in the form of a “Named Place”, “Distance” and “Direction” to obtain a geocode and transfer that to a file.

The example shown in fig.7 is of obtaining the latitude and longitude of a position “80 km SW of Toowoomba”, Queensland, Australia. The first step is to load the appropriate Gazetteer and select “Toowoomba” from it (**A**). There are a number of options, but I have selected the Toowoomba City (labelled POPL for Populated Place). The location of Toowoomba appears on the map in red (**B**). The distance “80” is typed into the Distance field and the pull down menus used to select “km” and “SSW” (**C**). The selected location appears on the map as a blue dot (**D**). The location, along with the latitude and longitude also appears on the bottom of the Gazetteer window (**E**). By right clicking on this area and selecting “Copy” that information can be copied and pasted into any Microsoft Windows compatible file (Word, Excel, Access, etc.). The Latitude and Longitude (to 1 arc-minute resolution) also appears (**F**), and this can similarly be copied to a file. Alternatively, by going to the Edit menu and select “Copy Lat/Long” the geocode can be copied to an accuracy of one arc-second.

One can also go to the map itself and zoom in to the point. Other coverages such as a road network (in ESRI Shape file format) can be loaded to allow more accurate positioning of the point – i.e. perhaps move it to the nearest road if collecting was done from a vehicle, etc. The selection tool can then be used to click on the point to obtain the geocode to one arc-second resolution. Again by right clicking with the mouse, or using Edit/Copy Lat/Long, that information can be copied to an appropriate file.

At present, the use of the eGaz program for areas outside Australia has a problem and the use of Distance and Direction from a point does not operate as it should. The rest of the program does work in these areas. The developers have been notified and it is hoped that this bug will be fixed shortly.

The BioLink Rapid Data Entry form (fig. 3) has a direct link to eGaz and the selection of a lat/long via eGaz will place that information back into the BioLink database (see above). Also, in reverse, it is easy to see where a specimen is on the eGaz map by clicking on the  button next to the Longitude.

Evaluation: eGaz is a powerful tool for museum and herbarium managers as mentioned in an earlier report (Chapman 2003). At present there is a major bug in the program, that reduces considerably it's use for Brazil and other areas outside Australia. As mentioned, this has been reported to the developers, and it is hoped that it will be fixed before long.

6.3.7 Modelling

BioLink includes two simple modelling algorithms, an early version of BIOCLIM (Nix, 1986, Busby 1991), and DOMAIN (Carpenter *et al.* 1993).

The program does allow some flexibility in the choosing of layers and in being able to include or exclude individual specimen records. It would have some application in data cleaning, however this is be limited and does not seem to be as valuable as other methods cited in Chapman (2003).

Evaluation: I have not attempted to use the implementations of these programs as included in BioLink, but they are tried and proven methodologies, and are simple to use. These are similar implementations to that of Diva-GIS (Hijmans *et al.* 2003) which was previously evaluated by me in an earlier report (Chapman 2003).

6.3.8 Tools

BioLink includes a number of tools for information and analysis and for managing pick lists, users, references, etc.. These include:

- Reference Manager
- Multimedia Manager
- Journal Manager
- Phrase (Glossary) Manager
- Trait and Note Types
- Electronic Gazetteer
- Mapping Assistant
- User Manager

- Specimen Loan Manager
- Predictive Distribution Modelling Tool
- Keyword

Evaluation: Most of these are self explanatory forms used for managing such things as user names and passwords, references for taxa etc. as well as other standard database administration tasks.

6.3.9 Reports

The Reports module contains a number of standard reports for such things as Sites for Taxa, Taxa for Sites, Label Manager, Data Entry Statistics, etc. as well as a simple SQL Query Builder. Pre-defined reports include:

- Sites for Taxa
- Taxa for Sites/Regions
- Taxa for Storage Location
- Material/Specimens For Taxon
- Taxon Statistics
- Taxonomic Checklist
- Specimen Labels
- Data Entry Statistics
- Report Editor
- Distribution Map
- Material for Lot/Bulk Sample/Trap

Evaluation: Without having data, it is difficult to assess the reports and their formats, however one would assume that they are standard SQL data reports.

7. Conclusion

BioLink is well-written and stable software. It is now in its 2nd full Version. This new version includes a Morphological Module which is a powerful addition. It also includes a character database based on the DELTA Standard. This is also a powerful addition, I wonder though if the two are best combined or kept separate. Other character analysis software, such as LUCID, have come a long way since the versions of DELTA included in this program, and are much more powerful and user friendly.

The program was written by Steve Shattuck and others at the CSIRO Division of Entomology in Canberra, Australia. As a result, it does have an entomological bias, but this does not preclude its use for other taxa. It probably has more relevance, however, to faunal databases than plant databases, but again, there is nothing to preclude its use for the latter.

Other than the problems mentioned with the use of eGaz outside Australia which the developers have been made aware of, and one minor problem with one help form coming up (also notified), the program seems stable and robust. It is public domain software available through the CSIRO Division of Entomology, in Canberra by emailing or writing to them (<http://www.biolink.csiro.au/>).

I would recommend it as suitable software for use as a *speciesLink*, and especially for any entomological database.

It has good import and export routines, including XML, which makes it ideal for distributed data exchange.

8. References:

- Busby, J.R. (1991). BIOCLIM – a bioclimatic analysis and prediction system. Pp. 4-68 in Margules, C.R. and Austin, M.P. (eds) *Nature Conservation: Cost Effective Biological Surveys and data Analysis*. Melbourne: CSIRO
- Chapman, A.D. (2003). *Notes on Environmental Data Quality – b. Data Cleaning Tools*. Internal report No.6 to CRIA, Draft in Preparation - June 2003.
- CSIRO (2001). BioLink. Canberra: CSIRO Division of Entomology.
<http://www.biolink.csiro.au/>.
- CSIRO (2003). BioLink 2.0 Build 459, RC 1, 21.5.03. Canberra: CSIRO Division of Entomology. CD-ROM.
- Dallwitz, M.J. (1980). A general system for coding taxonomic descriptions. *Taxon* **29**: 41-46.
- Dallwitz, M.J., Paine, T.A. and Zurcher, E.J. (1993). User's guide to the DELTA System: a general system for processing taxonomic descriptions. 4th edn.
<http://biodiversity.uno.edu/delta/>.
- Dallwitz, M.J., Paine, T.A. and Zurcher, E.J. (2000). Principles of interactive keys
<http://biodiversity.uno.edu/delta/>.
- Hijmans, R.J., Guarino, L., Bussink, C., Barrientes, I. and Rojas, E. (2003) *DIVA-GIS Version 3. A geographic information system for the analysis of biodiversity data*.
<http://www.diva-gis.org>
- Nix, H.A. (1986). A biogeographic analysis of Australian elapid snakes in Longmore, R.C. (ed). Atlas of Australian elapid snakes. *Australian Flora and Fauna Series* No. 7: 4-15. Canberra: Australian Government Publishing Service.
- Shattuck, S.O. (1997). eGaz, The Electronic Gazetteer. *ANIC News* **11**: 9
http://www.ento.csiro.au/research/natres/aninews/aninews11_09.html.
- Shattuck, S.O. and Fitzsimmons, N. (2000). *BioLink, The Biodiversity Information Management System*. Melbourne, Australia: CSIRO Publishing.
<http://www.biolink.csiro.au/>.
- Thiele, K.T., Rutter, G. and Yeates, D.K. (1997). LucID Interactive Key Software System Version 1. The Cooperative Research Centre for Tropical Pest Management, Brisbane.
- Watson, L. and Dallwitz, M.J. (1992 onwards) Grass Genera of the World: Descriptions, Illustrations, Identification, and Information Retrieval; including Synonyms, Morphology, Anatomy, Physiology, Phytochemistry, Cytology, Classification, Pathogens, World and Local Distribution, and References.
<http://biodiversity.uno.edu/delta/>.