

# **Atlas of Living Australia**

## **Review of online and desktop tools for the ALA**



**John Tann & Paul Flemons**

**February 2008**



## Introduction

This report presents a review of available software tools that could be used to support the Atlas of Living Australia.

These software tools include both desktop and online applications and searchable databases.

This is not an exhaustive list. Further tools can potentially be added through the [Atlas of Living Australia Tools Wiki](#).

In preparing this review we have investigated the following significant areas for appropriate tools:

- Data cleaning, validation and manipulation – eg spelling, misnaming, georeferencing, validation
- Visualisation – eg maps, graphs, images, tables etc
- Georeferencing – eg gazetteers
- Data analysis – eg environmental niche modelling; survey gap analysis
- Data Capture (of non-electronic data )– eg literature, digitisation of specimen data
- Taxonomy – eg identification
- Name resolution – eg name servers
- Provider interaction – eg building and preparing datasets, accepting feedback, communication
- Metadata – eg creation, collating, data discovery
- Environmental data – eg GIS layers, site specific data
- Bibliography – references
- Feedback – eg for errors, additions, alterations, quality
- User interface – personal or institutional, visual presentation or raw data, portable devices

Tools for database interaction were not investigated as part of the review, as these were considered to be intrinsic to the system architecture, a feature not yet determined. This included access protocols, data formats, metadata standards, and data exchange schema.

The software tools reviewed here will interact with the ALA in a variety of ways. There are examples of software tools that can be used as stand-alone applications, virtually independent of ALA architecture. Some tools lie firmly embedded inside other applications or websites, examples of potential for the ALA; or may possibly be able to be adapted specifically for the ALA. Many databases offer access to the public through a personal front-end. These databases may better benefit the ALA by a direct interaction, machine to machine. Other tools, such as desktop modelling tools, may use the ALA only as a vast source of data.

Comments for individual tools are encouraged. Personal experience, concerns, criticisms, difficulties, and comparisons can be extremely helpful. Comments on any of the tools can be added to the **Atlas of Living Australia Tools Wiki**: <http://alatools.pbwiki.com/>. **Invite Key:** *alatooledit*



# Table of Contents

<b>Introduction .....</b>	<b>iii</b>
Review format.....	1
<b>Data cleaning.....</b>	<b>3</b>
Data Tester.....	4
SALVIAS TaxonScrubber .....	7
speciesLink Data Cleaning.....	11
speciesLink spOutlier .....	14
<b>Mapping .....</b>	<b>16</b>
BerkeleyMapper.....	17
C-squares.....	20
Flash Earth.....	23
Gaia .....	25
Google Earth .....	27
Google Maps .....	31
InterMap .....	33
Live Search Maps / Virtual Earth.....	35
MapServer.....	37
NASA World Wind.....	41
Spatial Temporal Explorer.....	44
<b>Georeferencing.....</b>	<b>46</b>
BioGeomancer .....	47
Gazetteers .....	50
GEOLocate.....	56
MaNIS Georeferencing Calculator .....	58
infoXY .....	60
<b>Data analysis .....</b>	<b>62</b>

ANHAT – Australian National Heritage Assessment Tool .....	63
BIOCLIM – Bioclimatic prediction .....	65
Biodiverse.....	68
Biodiversity Statistics and Analysis .....	70
Biodiversity World.....	73
Biomapper.....	75
BioMaps .....	77
BRT – Boosted Regression Trees.....	80
BRUTO – Generalised Additive Modelling .....	82
CLUZ .....	84
DesktopGarp .....	86
DOMAIN .....	88
Eco-Tools .....	91
GBIF MAPA.....	94
GDM – Generalised Dissimilarity Modelling .....	96
Kepler .....	99
MARS – Multivariate Adaptive Regression Splines.....	102
Marxan .....	105
Maxent .....	108
openModeller .....	110
PRIMER – Multivariate Statistics for Ecologists .....	113
SGA tool – Survey Gap Analysis .....	115
SPRAT – Species Profile and Threats Database.....	118
Triana .....	120
Zonation .....	122
<b>Taxonomy .....</b>	<b>125</b>
Anatomical Atlas of Flies.....	126

APII – Australian Plant Image Index .....	128
APPD – Australian Plant Pest Database .....	131
AquaMaps .....	135
AVH – Australia’s Virtual Herbarium.....	138
DELTA - Intkey .....	141
FishBase .....	143
Fishes of Australia .....	147
LifeMapper .....	150
Lucid keys .....	153
Mandala .....	156
Morphbank .....	159
NBN Species Dictionary.....	162
PaDIL – Pest and Disease Image Library .....	165
PlantNET.....	168
<b>Name resolution.....</b>	<b>171</b>
AFD – Australian Faunal Directory .....	172
APNI – Australian Plant Names Index .....	174
Australian Insect Common Names.....	177
Australian Plant Census .....	179
Australian Plant Common Names .....	182
CAAB – Codes for Australian Aquatic Biota .....	184
Catalogue of Life .....	186
CAVS Biocodes – Census of Australian Vertebrate Species .....	189
ICTVdB – Virus Database.....	191
Index Fungorum .....	194
IPNI – International Plant Names Index.....	197
ITIS - Integrated Taxonomic Information System .....	199

LPSN – List of Prokaryotic names with Standing in Nomenclature.....	202
TROPICOS .....	204
uBio Taxonomic Name Server .....	206
WIN? – What’s Its Name? .....	210
<b>Genes .....</b>	<b>213</b>
BOLD-ID Barcode of Life Identification System.....	214
GenBank.....	217
Gene tools .....	220
Swami - The Next Generation Biology Workbench.....	223
<b>Metadata .....</b>	<b>225</b>
ASDD – Australian Spatial Data Directory .....	226
DIG – Discover Information Geographically.....	230
docBUILDER.....	233
EML - Ecological Metadata Language .....	235
MMI – Marine Metadata Interoperability .....	237
M3Cat.....	239
<b>Non-biological data .....</b>	<b>241</b>
ANRDL – Australian Natural Resources Data Library .....	242
DEM – Digital Elevation Model .....	244
GEBCO – General Bathymetric Chart of the Oceans.....	246
Map Maker.....	249
NDVI – Normalised Difference Vegetation Index .....	251
NVIS – National Vegetation Information System.....	254
OzClim .....	257
WorldClim .....	260
<b>Bibliography .....</b>	<b>262</b>
AnimalBase.....	263



Biodiversity Heritage Library.....	266
Botanicus.....	269
BUGZ – Bibliography of New Zealand Terrestrial Invertebrates.....	272
Google Scholar .....	274
PubMed.....	276
<b>Field Data Collection.....</b>	<b>279</b>
ArcPad .....	280
EFG – Electronic Field Guide .....	283
TDS Nomad .....	285
<b>Collaboration.....</b>	<b>288</b>
Scratchpads.....	289
<b>Other.....</b>	<b>292</b>
BioNet .....	293
DIVA-GIS.....	295
ESRI Explorers .....	298
GBIF RESTful Web Services .....	300
HerpNET .....	303
OZCAM – Online Zoological Collections of Australian Museums .....	305
Spotter – SPIRE Ontology Tool.....	307



## Review format

For each of the tools that have been reviewed here, the following template was used:

<b>Summary</b>	
Type of tool	eg application, framework with tools, language
Function	eg data manipulation, modelling, planning, metadata, viewing
Online / Desktop	
Computer infrastructure	Hardware, OS, software
Development status	eg active, experimental, beta; version, date
Time of use	
Licence	

**Description**

**Function**

- Data cleaning and manipulation
  - Data cleaning – spelling, misnaming
  - Data validating – taxonomy, geography
  - Georeferencing – applying latitude and longitude
- Visualisation tools
  - Maps
  - Images
- Analysis tools
  - Simple – distribution: single algorithm applied once
  - Complex – gap analysis: recursive
- Taxonomy
  - Identification tools, keys
- Provider interaction
  - Data preparation
  - Feedback
  - Communication
- Metadata
  - Creation
  - Data discovery
- Non-biological data
  - Environmental data
  - Site specific data
- Non-electronic data
  - literature
  - digitisation
- Feedback
- User interface
  - Personal or institutional use
  - Raw data or visual presentation

**Why use this tool?**

- What is the purpose of the tool?
- Are there alternatives?

**Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - Interest groups
  - General public
- ALA infrastructure
- Are there special skills required?

**How will the tool be used?**

- What are the data requirements?
- What ALA architecture is required to support the tool?
- What local architecture is required to support the tool?
- Is it used on the desktop or online?
- Is user input required?
- Can it run as a batch job?

**Where in the data chain could this tool be used?**

- Data source
- ALA central
- User's machine
- Pathways between these

**When could this tool be used?**

- Before data is made available to ALA
- As data is imported into ALA for storage
- While data is stored with ALA
- Whenever a specific event occurs
- At the time of a user request
- As a post process, after data is with the user

**Availability**

- Contact details
- Licence issues, intellectual property
- Cost
- Support
- Restrictions

**Comments**

Concerns, criticisms, difficulties, other features

## **Data cleaning**

## Data Tester

Summary	
Type of tool	Framework with tools
Function	Data cleaning and validation
Online / Desktop	Desktop
Computer infrastructure	Platform independent
Development status	Operational and expandable. Last update Oct 2006
Time of use	Pre-filter
Licence	Open Source

A set of tools to assist in checking the quality of biodiversity datasets.

### Description

A generic Java framework targeted to data cleaning and data validation. The idea behind this project has been originally conceived within the biodiversity informatics field. It followed the establishment of the first global networks that served primary data from biological collections. With the increase in the amount of shared data, which included researchers and policy makers among its users, data quality naturally gained importance. In this context, some networks started to develop tools and interfaces to help with data cleaning and data validation issues. The main idea of this project was to gather all knowledge from those first data cleaning tools and to produce a new framework that could serve as a common ground for implementing and running a large number of data tests.

The framework has been originally developed as open source software by the Reference Center on Environmental Information (CRIA) with funding from the Global Biodiversity Information Facility (GBIF) and the Gordon and Betty Moore Foundation. Despite being originated from the biodiversity informatics field, it is by no means bound or limited to this area. Its design pursued the following goals:

- To provide standard ways of interacting with the main components such as data tests, tests results, records and record sets, allowing different implementations for all of them.
- To be extensible and allow unlimited creation of new data tests that could be readily plugged into the framework.
- To be able to process record sets coming in different formats and from different sources (XML, relational database, etc).
- To allow the existence of parameterised data tests so that the same implementation could accept different configurations without the need of writing new tests.
- To make all data tests produce results in a standard format so that they can be handled programmatically.

Two Java packages were created: one containing the framework itself, and another containing a set of generic tests that can be useful in different situations.<sup>1</sup>

### Function

This is a suite of data cleaning and data validation tools.  
Tests that can be executed include the following:<sup>2</sup>

---

<sup>1</sup> <http://gbif.sourceforge.net/datatester/javadoc/>

<sup>2</sup> <http://www.gbif.org/Stories/STORY1128689677>

- Reporting unrecognized values for data elements (e.g. country names or basis of record values)
- Checking that coordinates fall within the boundaries of named geographic areas
- Finding scientific names that are not known to external lists such as the Catalogue of Life or nomenclators
- Checking that scientific names have an appropriate format
- Detecting numerical outliers

### **Why use this tool?**

Data quality is extremely important to both data users and data providers.

### **Who will use this tool?**

DataTester can be employed directly by data providers, other portals or persons preparing to perform analyses on data retrieved. In fact, the software is not limited to biodiversity data types, but those in fields other than biodiversity informatics can add tests for the kinds of errors that might be found in their data sets.<sup>3</sup>

### **How will the tool be used?**

The software is particularly suited to reporting on XML data sets, but can be applied to other data formats or relational databases. It allows programmers to develop new tests and to generalize tests so that they can work against multiple data standards (e.g. Darwin Core and ABCD schema). Each test may be associated with a severity (error, warning, info) to make it easier to focus on the most significant issues.<sup>4</sup>

Written in Java, this is a desktop application. The tester comes as four files, all platform independent:

- Framework
- Tests
- Source
- Documents

### **Where in the data chain could this tool be used?**

- Data source
- ALA central
- User's machine

### **When could this tool be used?**

- Before data is made available to ALA
- While data is stored with ALA
- As a post process, after data is with the user

### **Availability**

- Source: [https://sourceforge.net/project/showfiles.php?group\\_id=103853&package\\_id=165541](https://sourceforge.net/project/showfiles.php?group_id=103853&package_id=165541)
- News: <http://www.gbif.org/Stories/STORY1128689677>
- Documentation: <http://gbif.sourceforge.net/datatester/javadoc/>
- Licence: Open Source  
For conditions of use see SourceForge.net <https://sourceforge.net/tos/tos.php>

<sup>3</sup> <http://www.gbif.org/Stories/STORY1128689677>

<sup>4</sup> <http://www.gbif.org/Stories/STORY1128689677>

## Comments

GBIF have released three papers that discuss issues related to the quality of data:<sup>5</sup>

- **Principles and Methods of Data Cleaning** [http://www.gbif.org/prog/digit/data\\_quality/DataCleaning](http://www.gbif.org/prog/digit/data_quality/DataCleaning) (pdf)
- **Principles of Data Quality** [http://www.gbif.org/prog/digit/data\\_quality/DataQuality](http://www.gbif.org/prog/digit/data_quality/DataQuality) (pdf)
- **Uses of Primary Data** [http://www.gbif.org/prog/digit/data\\_quality/UsesPrimaryData](http://www.gbif.org/prog/digit/data_quality/UsesPrimaryData) (pdf)

---

<sup>5</sup> <http://www.gbif.org/Stories/STORY1124274724>



## SALVIAS TaxonScrubber

Summary	
Type of tool	Application
Function	Data cleaning
Online / Desktop	Desktop
Computer infrastructure	Widows, MS Access
Development status	Dated. Version 1.2 September 2004
Time of use	Data preparation. When data is imported into ALA
Licence	GNU General Public Licence

A stand alone tool for correction and standardization of spelling of plant species names,<sup>6</sup> and for detecting and flagging standard and non-standard species names.<sup>7</sup>

### Description

**SALVIAS TaxonScrubber** is a stand-alone application for automated standardization of taxonomic names. In addition to removing spelling errors in species names, TaxonScrubber splits concatenated information (such as *Genus + specific\_epithet + Author*) and stores each value in a separate field. This can be used to restructure flat-file specimen data prior to importing to a relational database. Although designed primarily for standardizing inventory data for the SALVIAS plots database, TaxonScrubber can be used whenever large numbers of taxonomic records need to be error-checked and reformatted.<sup>8</sup>

TaxonScrubber performs four basic actions:<sup>9</sup>

1. **Splitting of concatenated fields.** Epithets and authorities contained in single fields are split into separate fields. For example, the input string "*Quercus alba* L." is split into three fields, Genus = "*Quercus*", Species\_epithet = "*alba*", Sp\_auth = "L.". TaxonScrubber can split up to two subspecific levels off of a single name (e.g., *Quercus alba* var. *gunnisonii* Torr. fo. *Rugosa*).
2. **Recognition and removal of standard annotations.** TaxonScrubber contains an extensive library of Latin and English botanical annotations, their spelling variants, and abbreviations. Annotations such as "cf.", "aff.", "vel. sp. aff.", etc., are removed and stored in a separate field. Informal annotations of uncertainty, such as question marks, are treated as "cf." Any text not recognized as a standard annotation is stored in an additional annotation field, and flagged for inspection by the user.
3. **Standardization of spelling.** Once fields have been split, and extraneous text removed, TaxonScrubber matches names to a standard list of validly published names (currently, TaxonScrubber uses a world list of plant names; however, later releases of TaxonScrubber will have the option of loading name lists for other taxa). After flagging all names which match to the standard list, TaxonScrubber's "Hand scrub" utility provides pull-down menus for correcting remaining names to the standard world list. Names still unmatched at the end of the process can then be flagged as morphospecies names (e.g., *Miconia* sp.3), or as indets (e.g., *Miconia* sp.).
4. **Standardization of higher taxonomy.** TaxonScrubber standardizes all family names to match taxonomic concepts and spellings of the Missouri Botanical Garden's TROPICOS database. Future versions will allow the user to update higher taxonomy according to alternative

<sup>6</sup> <http://salvias.net/pages/whatissalvias.html>

<sup>7</sup> TaxonScrubber Ver 1.2 (September 2004)

<sup>8</sup> <http://salvias.net/pages/taxonscrubber.html>

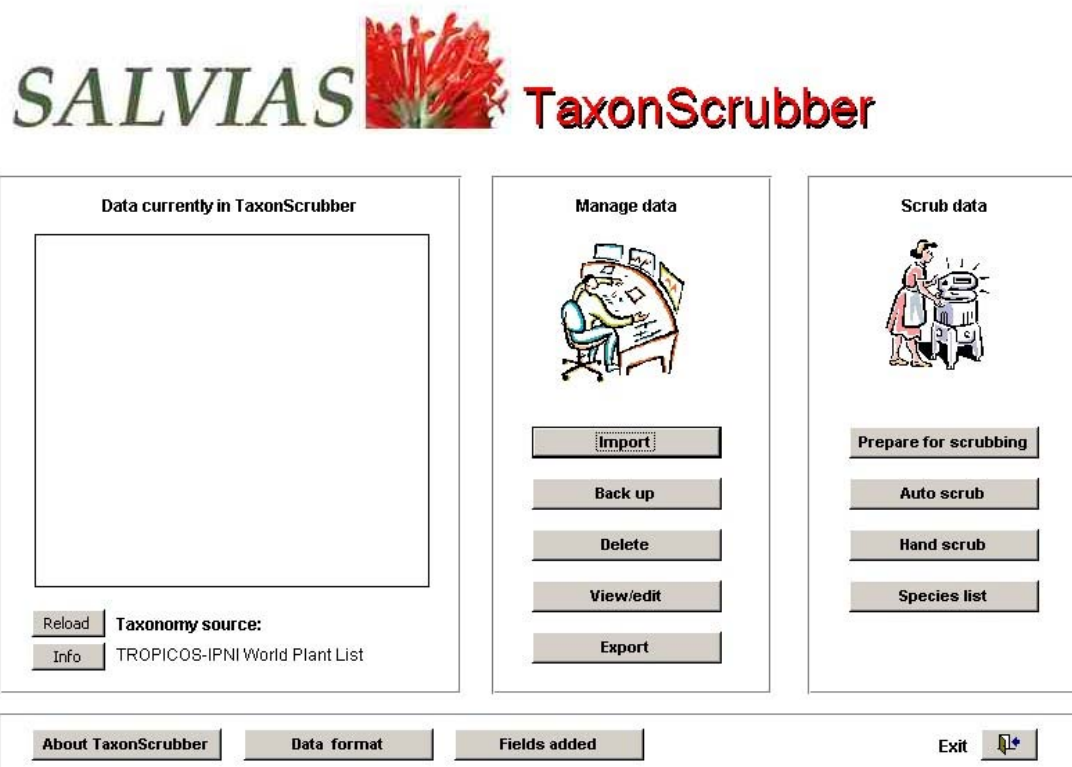
<sup>9</sup> <http://salvias.net/pages/taxonscrubber.html>

taxonomic concepts (for example, APG familial concepts; see [The Angiosperm Phylogeny Website](#)).

During the scrubbing process, TaxonScrubber generates new fields containing the results of the splitting and cleaning process, and various "flag fields" indicating the status of each name component (Family, genus, specific epithet, etc). These fields may be retained or deleted as needed upon export of the formatted the cleaned file.

Other TaxonScrubber features<sup>10</sup>

1. **File management.** TaxonScrubber imports, names, backs up, and manages source files within the database environment. Original files are left untouched until the user has completed the scrubbing process, and chooses to export the scrubbed file and replace the original.
2. **Archiving of source names.** Prior to scrubbing, TaxonScrubber archives the original names, unchanged, for comparison with the "scrubbed versions". After scrubbing, these fields can be deleted--or not--at the user's discretion.
3. **Hand-scrubbing.** TaxonScrubber features tools for manual inspection of taxonomic fields, including filters which display only records containing selected standard annotations, and matching to pull-down menus of standard names or names within the original file.



Screen shot of SALVIAS TaxonScrubber<sup>11</sup>

TaxonScrubber was developed by Brad Boyle in the Department of Ecology and Evolutionary Biology at University of Arizona, with support from the Center for Applied Biodiversity Science at Conservation International.<sup>12</sup>

<sup>10</sup> <http://salvias.net/pages/taxonscrubber.html>

<sup>11</sup> [http://salvias.net/pages/taxonscrubber\\_screenshot.html](http://salvias.net/pages/taxonscrubber_screenshot.html)

## Function

- Data cleaning and manipulation
  - Data cleaning – spelling, misnaming
  - File restructuring
- Taxonomy
- Provider interaction
  - Data preparation
- User interface
  - Personal use
  - Raw data

## Why use this tool?

- To correct and standardise the spelling of plant species names

## Who will use this tool?

- Data capture
- Data providers
  - Institutions
  - Private collections
- ALA infrastructure

## How will the tool be used?

Two files are required to run TaxonScrubber:

1. The main application – TaxonScrubber
2. Taxonomic database file – World plant list

**World plant list** is a lookup table for nearly 1 million plant names. Based on all names in a world list of vascular plant names from the Missouri Botanical Garden's TROPICOS database, with additional names of old world plants from the IPNI source databases. Compilation date: May 2003, reformatted for TaxonScrubber Ver. 1.2, Sept. 2004.<sup>13</sup>

- Windows, MS Access
- Desktop application
- User input is required

## Where in the data chain could this tool be used?

- Data source
- ALA central
- Pathways between these

## When could this tool be used?

- Before data is made available to ALA
- As data is imported into ALA for storage
- While data is stored with ALA

## Availability

- SALVIAS TaxonScrubber: <http://salvias.net/pages/taxonscrubber.html>
- Download TaxonScrubber: [http://salvias.net/Documents/TaxonScrubber\\_v12.zip](http://salvias.net/Documents/TaxonScrubber_v12.zip)
- Taxonomic database file: [http://salvias.net/Documents/TS\\_Taxon\\_Tables\\_worldlist\\_v12.zip](http://salvias.net/Documents/TS_Taxon_Tables_worldlist_v12.zip)

---

<sup>12</sup> TaxonScrubber Ver 1.2 (September 2004)

<sup>13</sup> <http://salvias.net/pages/taxonscrubber.html>

- Contact: Brad Boyle [bboyle@email.arizona.edu](mailto:bboyle@email.arizona.edu)
- Version 1.2. September 2004
- Cost: free
- Licence: GNU General Public Licence v3: <http://www.gnu.org/licenses/gpl.txt>).

#### Comments

- TaxonScrubber does not appear to be able to run as a batch job
- TaxonScrubber hasn't been updated for three years.

#### Q&A with Brad Boyle, TaxonScrubber creator, January 2008<sup>14</sup>

##### **Is TaxonScrubber still being maintained?**

*Yes and no. I originally developed it for my own use for cleaning data for import to SALVIAS. However, enough people were interested in it that I decided to make it available over our website. Although I have since issued a couple of updates, mostly bug-fixes, I will probably not be doing any further development, mostly because I do not want to continue working with Microsoft Access/Visual Basic. That said, I provide limited advice from time to time to people needing help with using the application.*

##### **If we were to use another Taxonomic database file, (for example a current extraction of the TROPICOS database, or a compilation of other species databases) is this possible/sensible?**

*Yes. The download would need to be reprocessed into the format which TaxonScrubber can read. Much would depend on the format of the original download, and whether or not it itself needs any cleaning. I would have to take a look at the list before committing to anything; if a lot of time would be involved, I would have to consider charging a consulting fee to compensate for time lost to other projects. Or, if you are familiar with programming in Access, you are welcome to try to produce a new reference database yourself.*

##### **Can TaxonScrubber be applied to other organisms eg fungi, animals?**

*Yes. Anything named with a Latin name. It's just a matter of having a taxonomic authority list. For example, I have used TaxonScrubber to check lists of North America birds against the AOU checklist.*

##### **Is there scope for someone (you/us?) to alter the program - for example to run as a batch job?**

*You're certainly welcome to try. If you are familiar with Visual Basic, the code is pretty transparent (if inelegant). However, although it does not run in command line, TaxonScrubber is still essentially a "batch" program, in the sense that it can process thousands of names at once.*

##### **Future directions**

*One of the reasons that I am no longer actively developing TaxonScrubber in its current form is that I would like to move it to a platform-independent Open Source version that would run as both a stand-alone and on the web. I haven't started work on this yet, but hope to release an initial version before the end of this year (2008). It's all a matter of finding the time, as usual.*

---

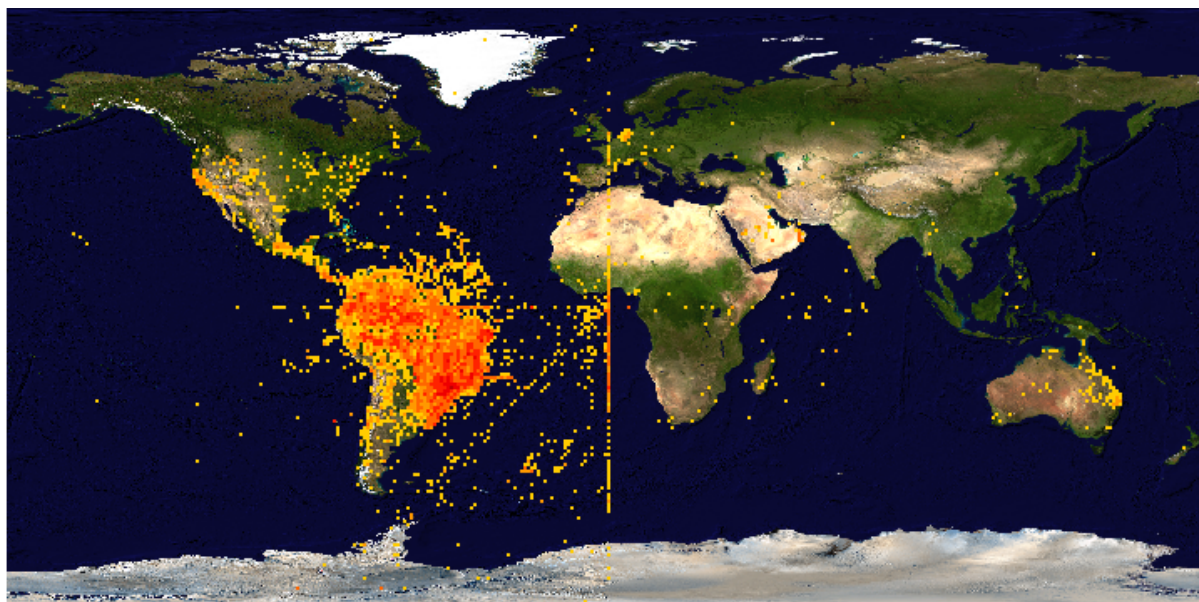
<sup>14</sup> Brad Boyle by email, 15 January 2008

## speciesLink Data Cleaning

Summary	
Type of tool	Set of tools
Function	Data cleaning
Online / Desktop	Online
Computer infrastructure	
Development status	Established
Time of use	Before the data is made available to ALA; while data is with the ALA
Licence	Negotiate to use locally

**Data cleaning** aims at helping curators in identifying possible errors and to standardize data. Records are not modified. The system just presents "suspect" records, recommending that they be checked by each author or curator.<sup>15</sup>

### Description



Geographic distribution of all records within the speciesLink network.<sup>16</sup> This map shows several sets of suspect data: data points on the Greenwich meridian probably have either a missing or zero longitude, similarly data points on the equator probably have either a missing or zero latitude; those data points on the line at a 45° angle to Greenwich meridian and the equator have the same value for both latitude and longitude; sea-based records concentrated on the southern side of this 45° line may have their latitude and longitude reversed; etc.

The *Data Cleaning* tool will summarise and report on:

- records without coordinates
- records in the sea
- repeated records/fields
- suspect taxonomy at family/genus/species/subspecies/author/duplicate
- suspect locality names of country/municipality

<sup>15</sup> <http://splink.cria.org.br/dc/index?criaLANG=en>

<sup>16</sup> <http://splink.cria.org.br/dc/index?criaLANG=en>

- suspect latitude and longitude
- outliers

### Function

- Data cleaning and manipulation
  - Data cleaning
  - Data validating – geography
- Visualisation tools
  - Maps
- User interface
  - Personal use
  - Data summary and visual presentation

### Why use this tool?

- To help curators identify data errors

### Who will use this tool?

- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Some skills are required

### How will the tool be used?

- Online tool when used for querying results of analysis
- User input is required
- Data is returned as a visual representation on a map, a summary report and data
- This tool is run on pre-loaded datasets, probably overnight<sup>17</sup>
- This tool should be modified for the ALA and run locally (see discussion below)
- *Data Cleaning* includes/links to the tools *spOutlier* and *infoXY*<sup>18</sup>

### Where in the data chain could this tool be used?

- Data source
- ALA central

### When could this tool be used?

- Before data is made available to ALA
- While data is stored with ALA

### Availability

- *Data Cleaning*: <http://splink.cria.org.br/dc/index?criaLANG=en>
- speciesLink: <http://splink.cria.org.br/index?criaLANG=en>
- Documentation: Environmental Data Quality - Data Cleaning Tools, Arthur Chapman 2004  
[http://splink.cria.org.br/docs/appendix\\_i.pdf](http://splink.cria.org.br/docs/appendix_i.pdf)
- Discussion list (in Portuguese): <http://www.cria.org.br/mailman/listinfo/splink-l>
- CRIA - Centro de Referência em Informação Ambiental, Brazil: <http://www.cria.org.br/>

<sup>17</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

<sup>18</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

- Licence: Negotiate with CRIA to modify and use *Data Cleaning* locally. CRIA is happy for others to use the code with acknowledgement<sup>19</sup>

#### Comments

- These are online tools for collections held by CRIA - Centro de Referência em Informação Ambiental, Brazil, and others.
- See also: Chapman, A.D. (2004). Environmental Data Quality – b. Data Cleaning Tools. Appendix I to */Sistema de Informação Distribuído para Coleções Biológicas: A Integração do Species Analyst e SinBiota. FAPESP/Biota process no. 2001/02175-5 March 2003 – March 2004./* Campinas, Brazil: CRIA 57 pp. [http://splink.cria.org.br/docs/appendix\\_i.pdf](http://splink.cria.org.br/docs/appendix_i.pdf).

Arthur Chapman, Australian Biodiversity Information Services has suggested that *Data Cleaning* is definitely the type of tool that the ALA needs and should use, and it would be best to obtain the code and run it or a modified version on the ALA. Although CRIA use some external datasets, they would probably not want the responsibility of running Australian data through the same tool, but that would need to be explored between the ALA and CRIA. It would be best to licence and use *Data Cleaning* for Australian collections.<sup>20</sup>

---

<sup>19</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

<sup>20</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

## speciesLink spOutlier

Summary	
Type of tool	Application
Function	Data cleaning
Online / Desktop	Online
Computer infrastructure	Browser, text entry or Excel spreadsheet file
Development status	Established
Time of use	Before data is made available to ALA, while data is with ALA
Licence	

**spOutlier** is an automated tool that detects outliers in latitude, longitude and altitude. It can also identify possible errors.<sup>21</sup>

### Description

Data is accepted as either an Excel spreadsheet, or can be input directly as comma or space delimited fields. Results are presented as a table with outliers highlighted, or optionally displayed on a map.

If the user indicates that the points are "marine" the system will highlight all points that occur inland,<sup>22</sup> and vice versa.



*Example data with outliers highlighted.*<sup>23</sup>

This is part of a suite of tools developed by CRIA as part of speciesLink, a distributed information system that integrates primary data from biological collections. The development was funded by FAPESP, GBIF, JRS Foundation and CRIA.<sup>24</sup>

### Function

- Data cleaning
  - Data cleaning
  - Data validating – geography
- Visualisation tools
  - Maps
- Feedback
- User interface

<sup>21</sup> <http://splink.cria.org.br/outlier?criaLANG=en>

<sup>22</sup> <http://splink.cria.org.br/outlier?criaLANG=en>

<sup>23</sup> <http://splink.cria.org.br/outlier>

<sup>24</sup> <http://splink.cria.org.br/index?criaLANG=en>



- Personal or institutional use
- Raw data and visual presentation

#### Why use this tool?

- Geographic error detection in collections

#### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
  - Casual users
- ALA infrastructure
- No special skills required

#### How will the tool be used?

- Comma or space delimited text, or an xls file
- Optional altitude
- Online tool
- User input is required
- *Data Cleaning* includes/links to this tool<sup>25</sup>

#### Where in the data chain could this tool be used?

- Data source
- ALA central

#### When could this tool be used?

- Before data is made available to ALA
- While data is stored with ALA

#### Availability

- spOutlier: <http://splink.cria.org.br/outlier>
- speciesLink: <http://splink.cria.org.br/index?criaLANG=en>
- CRIA - Centro de Referência em Informação Ambiental: <http://splink.cria.org.br>
- Licence: Negotiate with CRIA to modify and use *Data Cleaning* locally. CRIA is happy for others to use the code with acknowledgement<sup>26</sup>

#### Comments

Comment by Arthur Chapman:

*spOutlier* could be used directly via CRIA, however, I think it needs modification to use environmental outlier detection using my Reverse Jackknifing (as is done in the DIVA-GIS program and has been written for use in BioGeomancer) rather than as CRIA use it - i.e. only for geographic outliers using latitude, longitude and altitude as well as the offshore/onshore.<sup>27</sup>

<sup>25</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

<sup>26</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

<sup>27</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

## Mapping

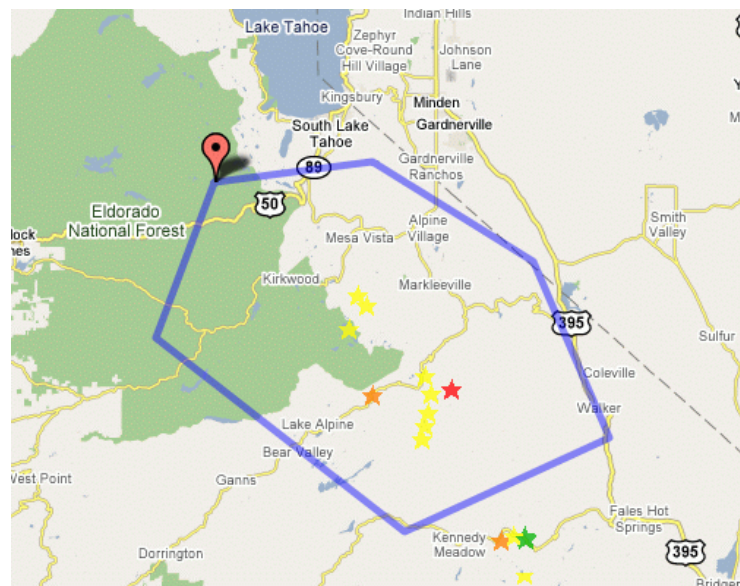
## BerkeleyMapper

Summary	
Type of tool	Web service
Function	Specimen mapping
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Experimental, under development
Time of use	At the time of a user request
Licence	Unknown

**BerkeleyMapper** displays point distribution maps from distributed sources.

### Description

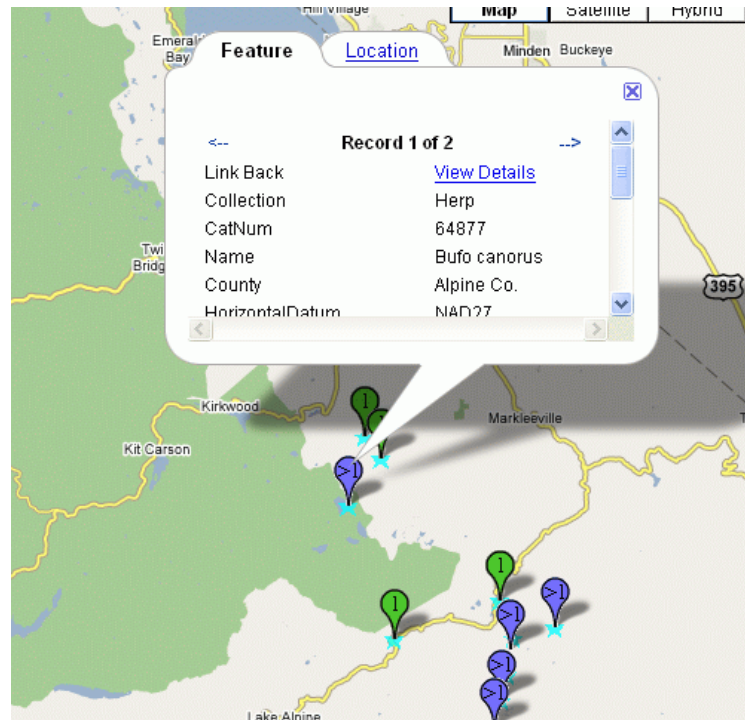
BerkeleyMapper was created and built using natural history museum specimen locations but is adaptable for mapping any collection of points. BerkeleyMapper is a collection of scripts and services bundled to provide integrated mapping functions using Google Maps, TerraServer, and open source mapping software components.<sup>28</sup>



*A user can select an area and extract the required sample data.*<sup>29</sup>

<sup>28</sup> <http://berkeleymapper.berkeley.edu/>

<sup>29</sup> <http://berkeleymapper.berkeley.edu/>



*Record details are available for any particular specimen.*<sup>30</sup>

BerkeleyMapper uses a Google Map base layer with its accompanying zooming and panning tools and, satellite and aerial photos; and overlays specimen data, DEM data, polygons and lines as a web service.

**Function**

- Visualisation tools
  - Maps
- User interface
  - Personal use
  - Raw data and visual presentation

**Why use this tool?**

This is a web service mapping tool in a convenient form. Very quickly users can map data, presenting their locations in a convenient and recognisable geospatial interface.

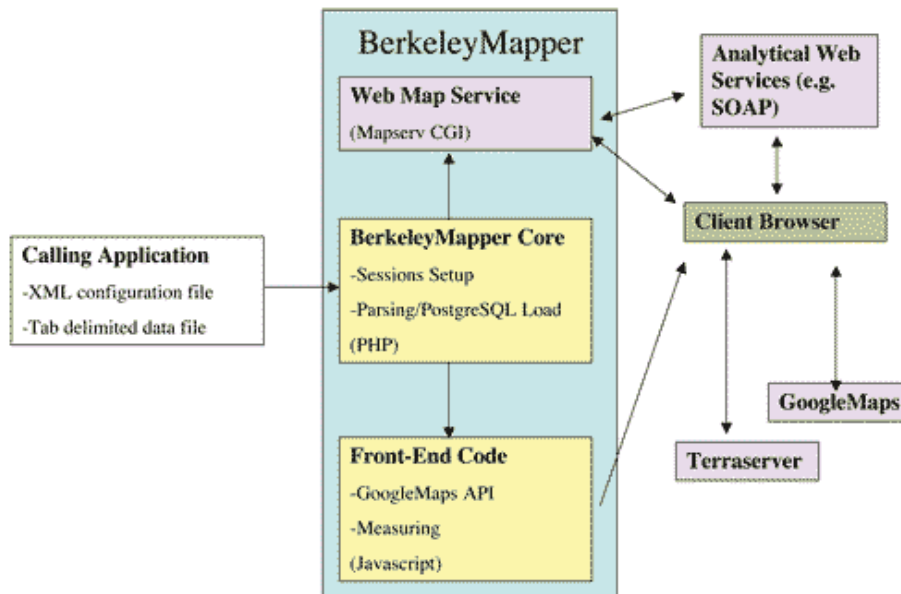
**Who will use this tool?**

- Data users
  - Expert
  - Interest groups
  - General public

**How will the tool be used?**

- Online web service
- User input required

<sup>30</sup> <http://berkeleymapper.berkeley.edu/>



*BerkeleyMapper operation.*<sup>31</sup>

Linking to BerkeleyMapper requires two steps:<sup>32</sup>

1. construct a URL from your web application that calls BerkeleyMapper with appropriate parameters.
2. point to an XML configuration file (that you store on your own server) that tells BerkeleyMapper how to behave. This file is referenced by the URL that you build in step #1, above.

[See Guide to Linking](#)

#### Where in the data chain could this tool be used?

- User's machine

#### When could this tool be used?

- At the time of a user request

#### Availability

- BerkeleyMapper: <http://berkeleymapper.berkeley.edu/>
- Guide to linking: <http://berkeleymapper.berkeley.edu/docs/help.html>
- Help: <http://berkeleymapper.berkeley.edu/docs/userhelp.php>
- Contact: [bnhm@berkeley.edu](mailto:bnhm@berkeley.edu)
- Licence: Unknown

#### Comments

As an example of an application of BerkeleyMapper see: Electronic Monograph of the Holarctic Engraver Beetles (Curculionidae: Scolytinae: Ipina)  
<http://peet.tamu.edu/projects/45/public/site/ipina/home/>

BerkeleyMapper has the ability to merge multiple occurrences at close locations, which separate as the map is zoomed to larger scales.

<sup>31</sup> <http://berkeleymapper.berkeley.edu/images/diagram.gif>

<sup>32</sup> <http://berkeleymapper.berkeley.edu/docs/help.html>

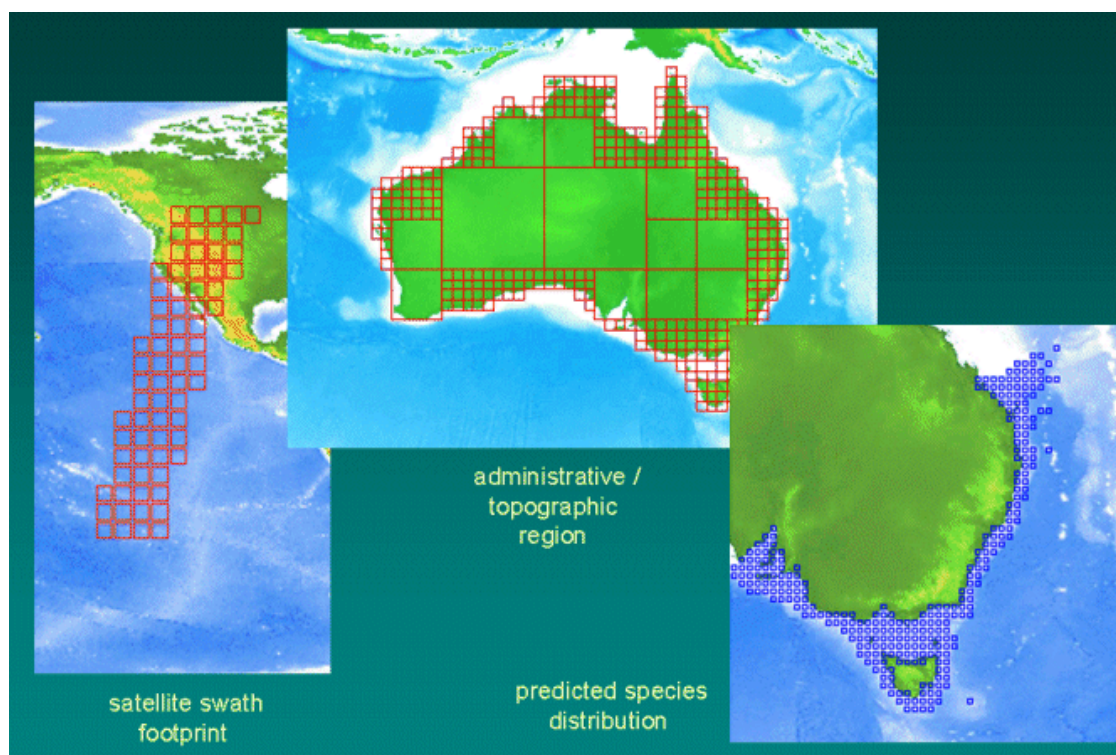
## C-squares

Summary	
Type of tool	Spatial data method, plus map tool
Function	Global locator system
Online / Desktop	C-squares mapper and converter are web based
Computer infrastructure	Tools require web server
Development status	Active, specification version 1.1 2005
Time of use	Data creation, data delivery, and when data is with user
Licence	Free to use without licence

**C-squares** (acronym for the **concise spatial query and representation system**) is a system of geocodes (actually a type of global grid) that provides a basis for simple spatial indexing of geographic features or data.<sup>33</sup>

### Description

The notation system of C-squares incorporates a compact encoding of latitude and longitude coordinates into a machine- and human-readable c-squares code, which can then be used either for spatial search or display via a suitable mapping application.<sup>34</sup>



Examples of dataset footprints displayed using C-squares.<sup>35</sup>

C-squares was devised as an improved (more precise) method for expressing dataset geographic extents in searchable metadata catalogues, in place of (or in addition to) conventional bounding rectangle representations. The method is also useful as a generic, interoperable notation for gridded data, for example a variety of datasets that describe environmental characteristics of

<sup>33</sup> <http://en.wikipedia.org/wiki/C-squares>

<sup>34</sup> <http://en.wikipedia.org/wiki/C-squares>

<sup>35</sup> <http://www.marine.csiro.au/csquares/about-csquares.htm>

global half-degree cells. C-squares can also simply be used for mapping (example: [CSIRO Marine and Atmospheric Research's "CAAB" application](#)), although the most value is obtained when the system is used for spatial search as well (e.g. [OBIS database](#), [CMAR's "MarLIN" metadatabase](#), etc.).

C-squares provides a hierarchical nomenclature for dividing 10°x10° [World Meteorological Organization \(WMO\) squares](#) into smaller units (each an individual "c-square") of 5°x5°, 1°x1°, 0.5°x0.5°, 0.1°x0.1°, etc., ... as fine as may be required. Each cell of the resulting subdivision is allocated a unique alphanumeric identifier (c-squares code), such that the position of an object or objects on the surface of the Globe can be represented by a set of one or more such codes that define the cell(s) within which the object occurs. Storing these codes as text identifiers, for example in a database, repository of spatial metadata, searchable text file or web page, then offers the functionality for a simple, text-based spatial search, without the requirement for any more complex geographic information system (GIS).<sup>36</sup>

The **CMAR c-squares mapper** is a perl utility which plots dataset extents (geographic footprints) on a range of base maps, according to a string of c-squares (and optional other parameters) passed to it via the web.<sup>37</sup>

The system was developed by Dr Tony Rees of CSIRO Marine and Atmospheric Research (CMAR) in 2001-2002 and is freely available for use worldwide without royalty or licence.

#### Function

- Visualisation tools
  - Maps

#### Why use this tool?

- To create searchable metadata
- For displaying a range of types of complex dataset footprints<sup>38</sup>

#### Who will use this tool?

- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
- Data users
  - Expert
  - Interest groups
- ALA infrastructure

#### How will the tool be used?

- Lat/long/c-squares converter is available online and as an Excel spreadsheet
- C-squares mapper is a web-accessible, perl based utility
- Installations of the c-squares mapper at CMAR currently run on UNIX and SUSE Linux and the mapper and Xmapper have also been successfully installed and run on Windows servers.<sup>39</sup>

---

<sup>36</sup> <http://en.wikipedia.org/wiki/C-squares>

<sup>37</sup> <http://www.marine.csiro.au/csquares/about-mapper.htm>

<sup>38</sup> <http://www.marine.csiro.au/csquares/about-csquares.htm>

<sup>39</sup> Installation notes – mapperV3.0 [http://sourceforge.net/project/showfiles.php?group\\_id=158386](http://sourceforge.net/project/showfiles.php?group_id=158386)

- Hardware requirements: web server capable of running cgi programs<sup>40</sup>
- Software requirements: Apache or equivalent web server, Perl, CGI for Perl<sup>41</sup>

#### **Where in the data chain could this tool be used?**

- The c-squares mapper and lat/long converter would sit on a server and be accessed through a user's machine

#### **When could this tool be used?**

- At the time of a user request
- As a post process, after data is with the user

#### **Availability**

- C-squares: <http://www.marine.csiro.au/csquares/>
- C-squares FAQ: <http://www.marine.csiro.au/csquares/csq-faq.htm>
- C-squares mapper code: [http://sourceforge.net/project/showfiles.php?group\\_id=158386](http://sourceforge.net/project/showfiles.php?group_id=158386)
- Lat/long converter: [http://sourceforge.net/project/showfiles.php?group\\_id=158386](http://sourceforge.net/project/showfiles.php?group_id=158386)
- Free to use. No royalty or licence.

#### **Comments**

C-squares use latitudes and longitudes as a reference, and so:

1. They are not square (In the southern hemisphere, the southern boundary will be shorter than the northern boundary.)
2. The size of a c-square will change with a change in latitude.  
See FAQ 16 [How big \(and what shape\) are individual c-squares on the ground?](#)

---

<sup>40</sup> For more detailed hardware requirements see Installation notes – mapperV3.0  
[http://sourceforge.net/project/showfiles.php?group\\_id=158386](http://sourceforge.net/project/showfiles.php?group_id=158386)

<sup>41</sup> <sup>41</sup> For more detailed softwarehardware requirements see Installation notes – mapperV3.0  
[http://sourceforge.net/project/showfiles.php?group\\_id=158386](http://sourceforge.net/project/showfiles.php?group_id=158386)



## Flash Earth

Summary	
Type of tool	Online application
Function	Mapping demonstration
Online / Desktop	Online
Computer infrastructure	Browser and Flash
Development status	Experimental for demonstration
Time of use	By developers for comparison
Licence	

**Flash Earth** is a mashup of several web mapping tools. It offers a visual comparison between several online mapping services.

### Description

This mapping demonstration is included as an example for developers, not as a working tool.

Flash Earth is an experimental application that uses satellite and aerial imagery from online mapping websites without official consent.<sup>42</sup>

This global mapping demonstration in Flash allows views using:<sup>43</sup>

- Google Maps
- Microsoft Virtual Earth (aerial)
- Microsoft Virtual Earth (labels)
- Yahoo! Maps
- Ask.com (aerial)
- Ask.com (physical)
- OpenLayers
- NASA Terra (daily)

There is a limited location search; a readout of latitude and longitude; and navigation tools.

This mashup is the creation of Paul Neave: [www.neave.com](http://www.neave.com).

### Function

- Visualisation tools
  - Maps

### Why use this tool?

This is a proof of concept of using Flash with multiple map overlays. It can be used to compare different online mapping services.

### Who will use this tool?

- Tool developers for demonstration

### How will the tool be used?

- Online tool

---

<sup>42</sup> <http://www.flashearth.com/>

<sup>43</sup> <http://www.flashearth.com/>

- User input is required

**Where in the data chain could this tool be used?**

- Not applicable

**When could this tool be used?**

- As a demonstration in the development phase of the ALA

**Availability**

- Flash Earth <http://www.flashearth.com/>
- Help: <http://www.neave.com/help/>
- Copyright: Paul Neave <http://www.neave.com/>
- Licence: Not available for commercial or private use<sup>44</sup>
- Cost: Free for online viewing

**Comments**

This is not a tool for use in practice, more an interesting Flash mapping application.

---

<sup>44</sup> <http://www.neave.com/help/>

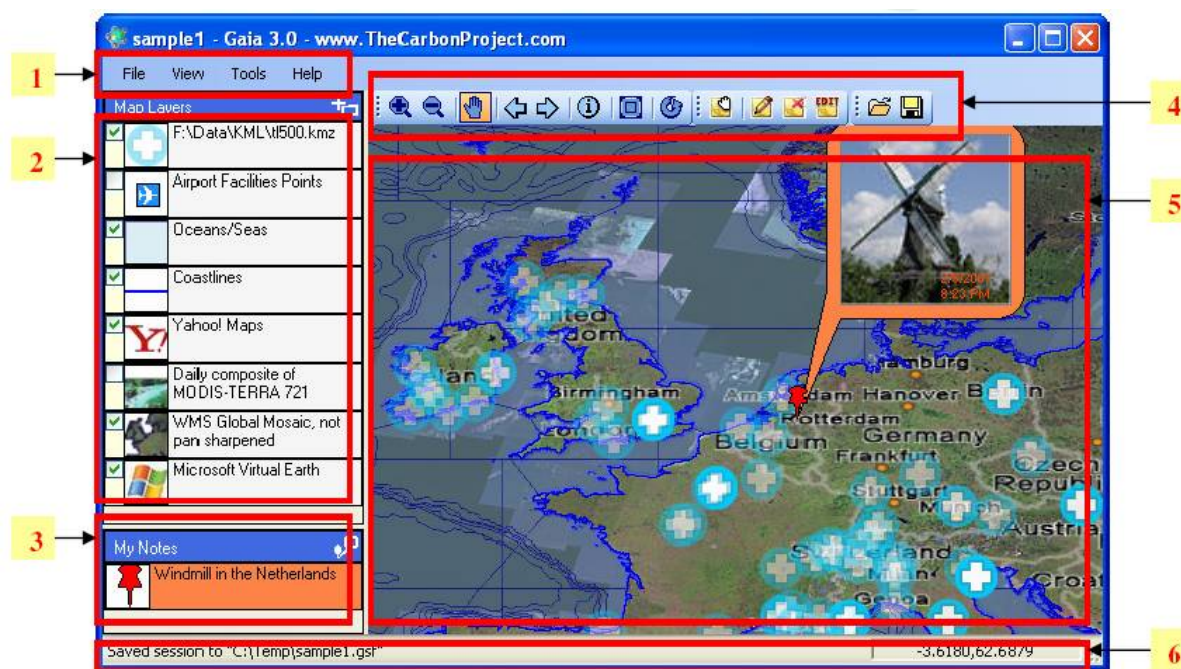
## Gaia

Summary	
Type of tool	Application
Function	Mapping
Online / Desktop	Desktop
Computer infrastructure	Windows .NET Framework
Development status	Version 3.1 Beta 2
Time of use	As a post process
Licence	Freeware

**Gaia 3** is a Windows application for accessing, visualizing and sharing location content.<sup>45</sup>

### Description

Gaia 3 lets you seamlessly access and use a vast array of location content and services from your Windows desktop – including Microsoft Virtual Earth, Yahoo! Maps, Google Earth KML/KMZ, OGC GML, ESRI Shapefiles, OGC WMS, WFS, WCS services and more.<sup>46</sup>



*Gaia main form has six general sections: 1. Main menu. 2. Map layers panel. 3. My notes panel. 4. Tools toolbar. 5. Map panel. 6. Status bar.*<sup>47</sup>

Gaia 3 was developed with CarbonTools PRO, an extension to Microsoft .NET Framework that supports advanced location content handling, mapping and sharing. Software developers can get the Gaia 3 source code by purchasing CarbonTools PRO.<sup>48</sup>

### Function

- Visualisation tools

<sup>45</sup> <http://www.thecarbonproject.com/gaia.php>

<sup>46</sup> <http://www.thecarbonproject.com/gaia.php>

<sup>47</sup> [http://www.thecarbonproject.com/pdfs/Gaia3\\_UserGuide.pdf](http://www.thecarbonproject.com/pdfs/Gaia3_UserGuide.pdf)

<sup>48</sup> <http://www.thecarbonproject.com/gaia.php>

- Maps
- Non-biological data
  - Environmental data
- User interface
  - Personal use
  - Visual presentation

#### **Why use this tool?**

Gaia uses geospatial content from different sources and overlay them into a single map view, with each layer individually configured and styled.<sup>49</sup>

#### **Who will use this tool?**

- Data users

#### **How will the tool be used?**

- Open Geospatial Consortium sources such as WMS, WCS, WFS<sup>50</sup>
- Commercial services such as Microsoft Virtual Earth, Yahoo Maps<sup>51</sup>
- Other formats: ESRI shapefiles, Google Earth KML/KMZ, DXF, MIF and GML<sup>52</sup>
- Windows .NET framework
- Desktop application

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- As a post process, after data is with the user

#### **Availability**

- Gaia: <http://www.thecarbonproject.com/gaia.php>
- Download: <http://www.thecarbonproject.com/gaia.php>
- Source Code and CarbonTools PRO : <http://www.thecarbonproject.com/dev.php>
- User's Guide: [http://www.thecarbonproject.com/pdfs/Gaia3\\_UserGuide.pdf](http://www.thecarbonproject.com/pdfs/Gaia3_UserGuide.pdf)
- Cost: free for Gaia viewer. CarbonTools PRO US\$1995
- Licence: Freeware

#### **Comments**

---

<sup>49</sup> [http://www.thecarbonproject.com/pdfs/Gaia3\\_UserGuide.pdf](http://www.thecarbonproject.com/pdfs/Gaia3_UserGuide.pdf)

<sup>50</sup> [http://www.thecarbonproject.com/pdfs/Gaia3\\_UserGuide.pdf](http://www.thecarbonproject.com/pdfs/Gaia3_UserGuide.pdf)

<sup>51</sup> [http://www.thecarbonproject.com/pdfs/Gaia3\\_UserGuide.pdf](http://www.thecarbonproject.com/pdfs/Gaia3_UserGuide.pdf)

<sup>52</sup> [http://www.thecarbonproject.com/pdfs/Gaia3\\_UserGuide.pdf](http://www.thecarbonproject.com/pdfs/Gaia3_UserGuide.pdf)

## Google Earth

Summary	
Type of tool	Application
Function	Global mapping
Online / Desktop	Desktop
Computer infrastructure	PC, Mac and Linux
Development status	Operational
Time of use	At data delivery, and as a post-process when data is with user
Licence	Four levels – personal use through to enterprise-wide

A search-based map tool for the creation and presentation of geographic and location-specific information.

### Description

**Google Earth** is a geographic browser - a powerful tool for viewing, creating and sharing interactive files containing highly visual location-specific information.<sup>53</sup> It maps the earth by superimposing images obtained by satellite and aerial photographs.

The viewer displays houses, the colour of cars, and even the shadows of people and street signs. The degree of resolution available is based somewhat on the points of interest, but most land (except for some islands) is covered in at least 15 meters of resolution, ... with the highest resolution at 15 cm. Google Earth allows users to search for addresses (for some countries only), enter coordinates, or simply use the mouse to browse to a location.<sup>54</sup>

Google Earth also uses digital elevation model (DEM) data, allowing a user to view the terrain in 3D.

There are several Google Earth products. See comparison table below.

#### Google Earth Free

It's the universe inside your PC; an atlas, encyclopedia and flight simulator, all rolled into one. Just point and zoom to any place you want to explore. Cities, mountains, and valleys are depicted in high-resolution 3D, along with related information.<sup>55</sup>

For personal use, Google Earth is available free.

#### Google Earth Plus

Intended for personal use, Google Earth Plus is a tool for mapping enthusiasts.

- Enhanced network access for faster performance
- Real-time GPS tracking and track/waypoint import -- for upload of data from select GPS devices
- Does not support export of tracks or waypoints to a GPS
- Greater-than-screen-resolution printing – for impressive hardcopies
- Spreadsheet importer – to import locations from .CSV files

Google Earth Plus is a \$20 annual subscription upgrade to Google Earth.

#### Google Earth Pro

For professional and commercial uses, Google Earth Pro uses the full power of the tool.

---

<sup>53</sup> <http://earth.google.com/>

<sup>54</sup> [http://en.wikipedia.org/wiki/Google\\_Earth](http://en.wikipedia.org/wiki/Google_Earth)

<sup>55</sup> <http://earth.google.com/earth.html>

- Enhanced printing, export high-resolution images up to 11" x 17" (4800 pixels)
  - Fastest data delivery speed
  - A research, presentation and collaboration tool for geo-specific information
  - 3D drawing tools
  - Import site plans, or other sites
  - Incorporate GIS data in file formats such as .shp and .tab
  - Transfer up to 2,500 locations by address or geospatial coordinates from a spreadsheet.
  - Optional premium data
- Google Earth Pro is a \$400 annual subscription

### Google Earth Enterprise

Google Earth Enterprise puts Google Earth inside the enterprise.

- It combines enterprise data with Google Earth data. Or an enterprise can host its own complete dataset.
- Any browser can view the Google Earth Enterprise implementation
- It has an enhanced search framework that enables integration of multiple search services through Java plug-ins, including the Google Search Appliance
- It has an area-based KML imagery data processing tool for creating super-overlays. ie a tool to create layers viewable on any Google Earth client and publishable to the public via any web server
- **Massively scalable** – hundreds of daily users can be supported from a single server, thousands from a small cluster
- **Unparalleled speed** –3D technology provides fluid access to remotely hosted multi-terabyte databases
- **Compatible with legacy GIS** – incorporate data in dozens of raster and point/vector GIS file formats

### Function

- Data testing tools
  - Data validating – geography
  - Georeferencing – applying latitude and longitude
- Analysis tools
  - Simple – distribution
- Visualisation tools
  - Maps

### Why use this tool?

- For presenting geographic based data on a dynamic map

### Who will use this tool?

- Data creation
- Data capture
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - General public
- ALA infrastructure

### How will the tool be used?

- Desktop version available for PC, Mac and Linux
- User input optional

### Where in the data chain could this tool be used?

- Data source
- ALA central
- User's machine

### When could this tool be used?

- Before data is made available to ALA
- At the time of a user request
- As a post process, after data is with the user

### Availability

- Download : <http://earth.google.com/>
- User guide: <http://earth.google.com/userguide/v4/>
- Version 4 2007
- Contact: <http://www.google.com/contact/>
- Cost depends on level:
  - Google Earth - free. For personal use
  - Google Earth Plus - US\$20 pa. Mapping and GPS support
  - Google Earth Pro - US\$400 pa. GIS, database support
  - Google Earth Enterprise: For complex integration with large datasets, servers
- Support for commercial systems
- Licence for personal non-commercial use: <http://earth.google.com/download-earth.html>
- Restrictions and permissions: <http://www.google.com/permissions/index.html>

### Comments

- Google Earth is a widespread and common global visualisation tool. It is able to incorporate user added content in the form of KML files.
- There are many similarities between the desktop Google Earth, and the online Google Maps.
- Google Earth Coordinate System and Projection
  - The internal coordinate system of Google Earth is geographic coordinates (latitude/longitude) on the World Geodetic System of 1984 (WGS84) datum.
  - Google Earth shows the earth as it looks from an elevated platform such as an airplane or orbiting satellite. The projection used to achieve this effect is called the General Perspective. This is similar to the Orthographic projection, except that the point of perspective is a finite (near earth) distance rather than an infinite (deep space) distance.<sup>56</sup>
- Only the Google Earth Enterprise version gives control over the base imagery.

---

<sup>56</sup> [http://en.wikipedia.org/wiki/Google\\_Earth](http://en.wikipedia.org/wiki/Google_Earth)

### Which version of Google Earth is right for you?

	Google Earth	Google Earth Plus	Google Earth Pro
License	Free version for home/personal use	Enhanced version for home/personal use	Professional version for commercial use
Price	Free	\$20*	\$400*
Imagery Database	Primary	Primary	Primary
Performance		Enhanced	Fastest
Fly to anywhere on the planet, or explore space	✓	✓	✓
Search for schools, parks, restaurants, and hotels	✓	✓	✓
Get driving directions	✓	✓	✓
Explore Featured Content	✓	✓	✓
Tilt and rotate the view in 3D	✓	✓	✓
Printing images	1000 pixels	1400 pixels	4800 pixels
Saving Images	1000 pixels	1000 pixels	4800 pixels
Drawing tools	✓	✓	✓
GPS data import (read only)**		✓	✓
Real-time GPS tracking		✓	✓
Spreadsheet data import		100 points	2500 points
Local Business Ads	✓	Optional	Optional
Support	Website only	Website, email (login issues only)	Website, email, chat
Measure area			✓
Movie Maker			✓
GIS data importing			✓
GDT traffic counts data***			\$200

\* subscription-based annual fee

\*\* verified support for Magellan and Garmin devices only

\*\*\* optional add-ons

from: [http://earth.google.com/product\\_comparison.html](http://earth.google.com/product_comparison.html)



# Google Maps

Summary	
Type of tool	Application
Function	Global mapping
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Operational
Time of use	At data delivery, and as a post-process when data is with user
Licence	

A browser-based map service.

## Description



**Google Maps** allows you to view basic or custom maps and local business information. You can zoom and pan map data and integrated satellite images.

Depending on your location, you can create personalized, annotated, customized maps using Google Maps. Your maps can contain the following:<sup>57</sup>

- Placemarks
- Lines
- Shapes

Once you have created a map, you can:<sup>58</sup>

- Add descriptive text, including rich text and HTML
- Embed photos and videos in your map
- Share your maps with others
- Open it in Google Earth

Google maps can be viewed in several ways:

- At Google Maps website
- Through a web page with an embedded Google Map
- On a mobile phone or other device

<sup>57</sup> <http://maps.google.com/support/bin/answer.py?answer=68480>

<sup>58</sup> <http://maps.google.com/support/bin/answer.py?answer=68480>

- As a private map created through Google Earth Enterprise

#### Function

- Analysis tools
  - Simple – geospatial
- Visualisation tools
  - Maps

#### Why use this tool?

- To geographically represent data in point, line or area form on a map.  
See also Google Earth

#### Who will use this tool?

- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - Interest groups
  - General public

#### How will the tool be used?

- Online tool used within a browser
- File format is KML
- User input not required

#### Where in the data chain could this tool be used?

- User's machine

#### When could this tool be used?

- Before data is made available to ALA, for creating maps
- At the time of a user request, for displaying maps

#### Availability

- Google maps: <http://maps.google.com.au/>
- User guide: <http://maps.google.com/support/bin/answer.py?answer=68259>
- Support: <http://maps.google.com/support/>
- Google Maps API for embedding in web pages: <http://www.google.com/apis/maps/>
- Licence: Personal or internal business use; or through a website using an API.
- Terms of Service: [http://www.google.com/intl/en\\_au/help/terms\\_maps.html](http://www.google.com/intl/en_au/help/terms_maps.html)
- API Terms of Service: <http://code.google.com/apis/maps/terms.html>
- Cost: Free

#### Comments

- The highest zoom gives a scale of about 1:1000 (2 cm : 20 m)
- We have no control over the base imagery

## InterMap

Summary	
Type of tool	Application
Function	Web mapping
Online / Desktop	Online
Computer infrastructure	OS independent
Development status	Sleepy development. Version 2.1 alpha 1, March 2007
Time of use	At time of user request
Licence	GNU General Public Licence

**InterMap** is an Internet mapping application that allows the user to combine interactive maps from distributed Internet Map Servers in a browser.<sup>59</sup>

### Description

InterMap supports OpenGIS WMS and ESRI-ArcIMS and can be fully integrated with the GeoNetwork Metadata portal.<sup>60</sup>

The targeted user community is people working in developing countries with poor internet connections.<sup>61</sup>

### Function

- Visualisation tools
  - Maps
- Non-biological data
  - Environmental data
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

- Mapping over the internet

### Who will use this tool?

- Data users

### How will the tool be used?

- Platform independent
- Requires JDK - Java Development Kit
- InterMap installs a servlet accessible through the browser
- Online application

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

---

<sup>59</sup> <http://sourceforge.net/projects/intermap>

<sup>60</sup> <http://sourceforge.net/projects/intermap>

<sup>61</sup> [http://sourceforge.net/forum/forum.php?thread\\_id=1238849&forum\\_id=266961](http://sourceforge.net/forum/forum.php?thread_id=1238849&forum_id=266961)

- At the time of a user request

**Availability**

- InterMap: <http://sourceforge.net/projects/intermap/>
- Download: [http://sourceforge.net/project/showfiles.php?group\\_id=78244](http://sourceforge.net/project/showfiles.php?group_id=78244)
- Version 2.1 alpha 1, March 2007
- Licence: GNU General Public Licence
- Limited support

**Comments**

This program has a quiet presence on SourceForge. A handful of requests over the past 3 years, and less than 100 downloads per month. No recent forum posts.

## Live Search Maps / Virtual Earth

Summary	
Type of tool	Application
Function	Global mapping
Online / Desktop	Desktop
Computer infrastructure	Web browser
Development status	Active development. Version 6.0
Time of use	At data delivery
Licence	Microsoft

**Live Search Maps** is a web mapping service provided as a part of Microsoft's Windows Live online applications services suite and powered by Microsoft's **Virtual Earth**.<sup>62</sup>

### Description

Live Search Maps has detailed street maps for many cities throughout the world. It also has satellite images, with an image resolution that varies across the globe.

Live Search Maps is the public face of the Virtual Earth platform.

Virtual Earth provides powerful mapping capabilities for integrating location information into business solutions. With Virtual Earth, organizations can create applications that enable users to visually interact with complex information. Additionally, its service-oriented architecture enables companies to develop innovative solutions that take advantage of customizable features and imagery, dynamic maps, driving directions, and powerful data visualization and reporting capabilities without significant financial investments.<sup>63</sup>

### Function

- Visualisation tools
  - Maps
- User interface
  - Personal

### Why use this tool?

- To geographically present or view data on a map.

### Who will use this tool?

- Data capture
  - Curators
- Data providers
  - Institutions
- Data users
  - Expert
  - Interest groups
  - General public

### How will the tool be used?

- This is an online tool. Does not work well with some browsers

---

<sup>62</sup> [http://en.wikipedia.org/wiki/Live\\_Search\\_Maps](http://en.wikipedia.org/wiki/Live_Search_Maps)

<sup>63</sup> <http://www.microsoft.com/virtualearth/platform/>

**Where in the data chain could this tool be used?**

- Data source
- User's machine

**When could this tool be used?**

- Before data is made available to ALA
- As a post process, after data is with the user

**Availability**

- Live Search Maps: <http://maps.live.com/>
- Live Search Maps developer site: <http://www.viavirtualearth.com/vve/Default.ashx>
- MSDN Developer Centre: <http://msdn2.microsoft.com/en-us/virtualearth/default.aspx>
- Licence: Microsoft <http://www.microsoft.com/virtualearth/product/licensing.aspx>
- Cost: depends on the specific licence option
- Code of Conduct: <http://maps.live.com/Help/en-us/CodeOfConduct.htm>

**Comments**

This seems better suited to commercial and business use.

A visual comparison of images from Virtual Earth and other mapping software is available at: Flash Earth <http://www.flashearth.com/>.

# MapServer

Summary	
Type of tool	Application
Function	Serving maps for the web
Online / Desktop	Server with online map viewing
Computer infrastructure	Windows, Mac OS, Linux, Solaris. Requires a web server
Development status	Active and well developed. Version 5.0 September 2007
Time of use	Data creation. At time of user request
Licence	Open Source

**MapServer** is a web-based mapping tool.

## Description

MapServer is an Open Source development environment for building spatially-enabled internet applications. MapServer is not a full-featured GIS system, nor does it aspire to be. Instead, MapServer excels at rendering spatial data (maps, images, and vector data) for the web.

Beyond browsing GIS data, MapServer allows you create "geographic image maps", that is, maps that can direct users to content.<sup>64</sup>

MapServer supports Open Geospatial Consortium (OGC) standards, including Web Map Service (WMS) and Web Feature Service (WFS). MapServer works with PostgreSQL and its PostGIS extension, and supports proprietary GIS formats including ESRI's Shapefile format.<sup>65</sup>

In its most basic form, MapServer is a CGI program that sits inactive on your Web server. When a request is sent to MapServer, it uses information passed in the request URL and the Map File to create an image of the requested map. The request may also return images for legends, scale bars, reference maps, and values passed as CGI variables.<sup>66</sup>

MapServer can be greatly extended and customized. It can be built to support many different input data formats and output types. This is done at the time the MapServer binary is compiled. See the MapServer Home Page for a full list of current features. Many of the features that are not 'built-in', are enabled through the use of OGR, a library of tools giving access to GIS file formats.<sup>67</sup>

**MapScript** provides a scripting interface for MapServer for the construction of Web and stand-alone applications. MapScript is used independently of CGI MapServer, it is a loadable module that adds MapServer capability to your favourite scripting language. MapScript currently exists in PHP, Perl, Python, Ruby, Tcl, Java, and C# flavours.<sup>68</sup>

---

<sup>64</sup> <http://mapserver.gis.umn.edu/>

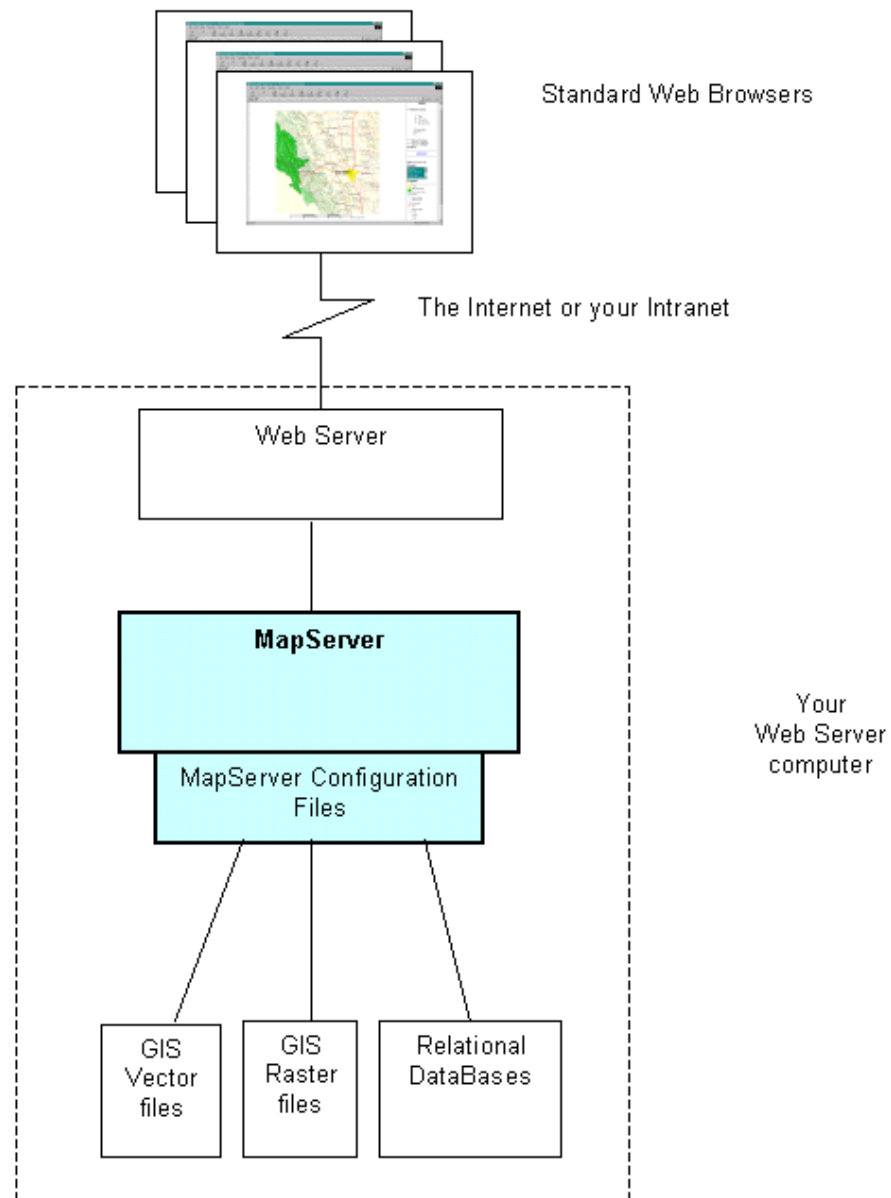
<sup>65</sup> <http://en.wikipedia.org/wiki/Mapserver>

<sup>66</sup> [http://mapserver.gis.umn.edu/new\\_users/](http://mapserver.gis.umn.edu/new_users/)

<sup>67</sup> [http://mapserver.gis.umn.edu/new\\_users/](http://mapserver.gis.umn.edu/new_users/)

<sup>68</sup> [http://mapserver.gis.umn.edu/new\\_users/](http://mapserver.gis.umn.edu/new_users/)

## MapServer Architectural Overview



*A conceptual diagram of a typical **MapServer** application.<sup>69</sup>*

### Features<sup>70</sup>

- Advanced cartographic output
  - Scale dependent feature drawing and application execution
  - Feature labelling including label collision mediation
  - Fully customizable, template driven output
  - TrueType fonts
  - Map element automation (scalebar, reference map, and legend)
  - Thematic mapping using logical- or regular expression-based classes
- Support for popular scripting and development environments
- PHP, Python, Perl, Ruby, Java, and C#
- Cross-platform support

<sup>69</sup> [http://mapserver.gis.umn.edu/new\\_users/msappdiagram/image\\_view\\_fullscreen](http://mapserver.gis.umn.edu/new_users/msappdiagram/image_view_fullscreen)

<sup>70</sup> <http://mapserver.gis.umn.edu/>



- Linux, Windows, Mac OS X, Solaris, and more
- A multitude of raster and vector data formats
  - TIFF/GeoTIFF, EPPL7, and many others via GDAL
  - ESRI shapefiles, PostGIS, ESRI ArcSDE, Oracle Spatial, MySQL and many others via OGR
  - Open Geospatial Consortium (OGC) web specifications
    - WMS (client/server), non-transactional WFS (client/server), WMC, WCS, Filter Encoding, SLD, GML, SOS
- Map projection support

MapServer was originally developed by the University of Minnesota (UMN) ForNet project in cooperation with NASA and the Minnesota Department of Natural Resources (MNDNR). Presently, the MapServer project is hosted by the TerraSIP project, a NASA sponsored project between the UMN and consortium of land management interests.

The software is maintained by a growing number of developers (nearing 20) from around the world and is supported by a diverse group of organizations that fund enhancements and maintenance.<sup>71</sup>

#### Function

- Visualisation tools
  - Maps
- User interface
  - Visual presentation

#### Why use this tool?

- To deliver maps to the desktop

#### Who will use this tool?

- ALA infrastructure

#### Skills

In addition to learning how the different components of a MapServer application work together and learning Map File syntax, building a basic application requires some conceptual understanding and proficiency in several skill areas.

You need to be able to create or at least modify HTML pages and understand how HTML forms work. Since the primary purpose of a MapServer application is to create maps, you will also need to understand the basics of geographic data and likely, map projections. As your applications get more complex, skills in SQL, DHTML/JavaScript, Java, databases, expressions, compiling, and scripting may be very useful.<sup>72</sup>

#### How will the tool be used?

- MapServer is available for Windows, Linux, and Mac OS, Solaris and more.
- Source code is available
- Server application
- An HTTP web server is required

---

<sup>71</sup> <http://mapserver.gis.umn.edu/>

<sup>72</sup> [http://mapserver.gis.umn.edu/new\\_users/](http://mapserver.gis.umn.edu/new_users/)

**Where in the data chain could this tool be used?**

- ALA central

**When could this tool be used?**

- At the time of a user request

**Availability**

- MapServer: <http://mapserver.gis.umn.edu/>
- Download: <http://mapserver.gis.umn.edu/download/current>
- Documentation: <http://mapserver.gis.umn.edu/docs> (extensive)
- Community: <http://mapserver.gis.umn.edu/community>
- Licence: no restrictions <http://mapserver.gis.umn.edu/License>
- Some portions are copyright State of Minnesota, Land Management Information Center<sup>73</sup>
- Cost: free
- Open source. Free to use, copy, redistribute, modify, sell.
- No warranty

**Comments**

---

<sup>73</sup> <http://mapserver.gis.umn.edu/License>

## NASA World Wind

Summary	
Type of tool	Application
Function	Global mapping
Online / Desktop	Desktop
Computer infrastructure	Windows
Development status	Active development. Version 1.4 February 2007
Time of use	At data delivery, and as a post-process when data is with user
Licence	NASA Open Source

**World Wind** is open source Windows software that lets you zoom from satellite altitude into any place on Earth. Leveraging Landsat satellite imagery and Shuttle Radar Topography Mission data, World Wind lets you experience Earth terrain in visually rich 3D, just as if you were really there.<sup>74</sup>

### Description



World Wind is a free open source virtual globe developed by NASA and the open source community for use on personal computers running Microsoft Windows. The program overlays NASA and USGS satellite imagery, aerial photography, topographic maps and publicly available GIS data on 3D models of the Earth.

Users interact with the world by rotating it, tilting the view, and zooming in and out. Five million placenames, political boundaries, latitude/longitude lines, and other location criteria can be displayed. World Wind provides the ability to browse maps and geospatial data on the internet using the OGC's WMS servers and WFS for downloading placenames, import ESRI shapefiles and kml/kmz files.

The resolution is at least 15 meters per pixel.

World Wind uses digital elevation model (DEM) data collected by NASA's Shuttle Radar Topography Mission. This means one can view topographic features such as the Grand Canyon or Mount Everest in three dimensions. In addition, WW has bathymetry data which allows users to see ocean features, such as trenches and ridges, in 3D.<sup>76</sup>

<sup>74</sup> <http://worldwind.arc.nasa.gov/>

<sup>75</sup> <http://worldwind.arc.nasa.gov/>

<sup>76</sup> [http://en.wikipedia.org/wiki/World\\_Wind](http://en.wikipedia.org/wiki/World_Wind)

## **Datasets available<sup>77</sup>**

### **Image/terrain datasets:**

- Blue Marble Next Generation imagery
- Landsat 7 imagery
  - NLT Landsat (Visible & Pseudo Colour)
  - Geocover 1990 & 2000 (pseudo; 1990 layer was produced from Landsat 4 & 5 images)
  - OnEarth (visible & pseudo)
  - i-cubed (visible)
- Zoomit! imagery (community produced layer)
  - LINZ[2] (montaged colour aerial photography of New Zealand)
  - GSWA[3] (Topographic and geological maps of Western Australia)
- SRTM (SRTM30Plus[4]/SRTMv2/USGS NED) terrain data (includes bathymetry)

### **Animated data layers:**

- Scientific Visualization Studio
- MODIS
- GLOBE
- NRL Real-Time Weather

### **Function**

- Visualisation tools
  - Maps

### **Why use this tool?**

- To present geographically based scientific information on a dynamic image of the world
- World Wind has a more scientific function than Google Earth

### **Who will use this tool?**

- Data users
  - Expert
  - Interest groups
  - General public

### **How will the tool be used?**

- Uses online data, and requires a high bandwidth
- Desktop application on Windows
- User input required

### **Where in the data chain could this tool be used?**

- User's machine

### **When could this tool be used?**

- At the time of a user request
- As a post process, after data is with the user

### **Availability**

- NASA World Wind: <http://worldwind.arc.nasa.gov/>
- World Wind FAQ: [http://www.worldwindcentral.com/wiki/World\\_Wind\\_FAQ](http://www.worldwindcentral.com/wiki/World_Wind_FAQ)

---

<sup>77</sup> [http://en.wikipedia.org/wiki/World\\_Wind](http://en.wikipedia.org/wiki/World_Wind)

- NASA Open Source: <http://worldwind.arc.nasa.gov/worldwind-nosa-1.3.html>
- Some images are copyright:  
[http://www.worldwindcentral.com/wiki/Copyright\\_Questions\\_When\\_Using\\_Images\\_From\\_World\\_Wind](http://www.worldwindcentral.com/wiki/Copyright_Questions_When_Using_Images_From_World_Wind)
- Cost: Free
- Data can be freely modified, redistributed and used on web sites.<sup>78</sup>

### Comments

This is a geographic viewer with similarities and comparisons with Google Earth.

#### **Which is better, Google Earth or World Wind?**

Both *NASA World Wind* and *Google Earth* are sophisticated programs with amazing feature sets, and they both have features that the other would benefit from. Despite how similar both programs seem they are not entirely comparable due to significantly different focuses, development history and philosophies. Many people use both programs depending on what they wish to accomplish.<sup>79</sup>

---

<sup>78</sup> [http://en.wikipedia.org/wiki/World\\_Wind](http://en.wikipedia.org/wiki/World_Wind)

<sup>79</sup> [http://www.worldwindcentral.com/wiki/World\\_Wind\\_FAQ#Installation](http://www.worldwindcentral.com/wiki/World_Wind_FAQ#Installation)

## Spatial Temporal Explorer

Summary	
Type of tool	Application
Function	Viewing metadata, maps, other data
Online / Desktop	Desktop
Computer infrastructure	Data held by GCMD
Development status	Last modified April 2006
Time of use	Post process
Licence	

This application is designed for viewing metadata, maps and other data that contains temporal and/or spatial coverages<sup>80</sup> of the Antarctic region.

### Description

Software for an interactive spatial and temporal map of specific Antarctic data covering 1945 to present and from longitudes 50 to 170 east. See [SpatialTemporalExplorer](#) for live application. Contains source code, Flash MX library, actionscript code and several ColdFusion CFCs for Flash remoting. Authored using Flash MX2004 version 7.2.<sup>81</sup>

### Function

- Visualisation tool
  - Maps

### Why use this tool?

- To view available data in the Antarctic region

### Who will use this tool?

- Data users
  - General public

### How will the tool be used?

- Desktop application
- Requires Flash player

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- At the time of a user request
- As a post process, after data is with the user

### Availability

- Source: [http://aadc-maps.aad.gov.au/aadc/portal/download\\_file.cfm?file\\_id=1502](http://aadc-maps.aad.gov.au/aadc/portal/download_file.cfm?file_id=1502)
- No metadata available
- Licence not mentioned
- Free to download upon registration

<sup>80</sup> <http://aadc-maps.aad.gov.au/aadc/search/SpatialTemporalExplorer.cfm>

<sup>81</sup> [http://aadc-maps.aad.gov.au/aadc/portal/download\\_file.cfm?file\\_id=1502](http://aadc-maps.aad.gov.au/aadc/portal/download_file.cfm?file_id=1502)

- Support unknown
- Date revised 17 April 2006

**Comments**

The downloaded files seemed to be incomplete.

## Georeferencing



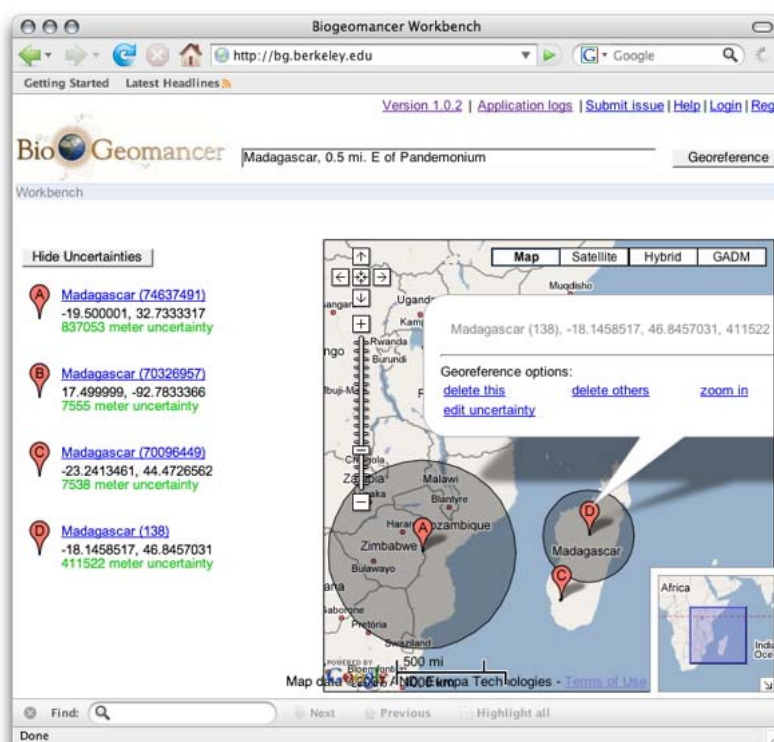
## BioGeomancer

Summary	
Type of tool	Application
Function	Georeferencing
Online / Desktop	Online. Single reference and batch mode
Computer infrastructure	Web browser
Development status	Active development. Version 1.1.0 October 2007
Time of use	In preparing data, while checking data, and at time of user request
Licence	Unknown

**BioGeomancer** workbench is an online georeferencing tool that returns the geographic coordinates of a location submitted as plain text.

### Description

As an input, BioGeomancer can accept plain English descriptions, such as “15 km SW Bathurst”, convert these to computer-readable geographic locations and will display these locations on a Google map.



*BioGeomancer*<sup>82</sup>

BioGeomancer implements the point-radius method. This method returns a single point and a circle of "uncertainty" around that point. Uncertainty depends on the quality of the locality description and the quality of the data used to map it.<sup>83</sup>

<sup>82</sup> <http://code.google.com/p/biogeomancer-workbench/>

<sup>83</sup> <http://groups.google.com/group/biogeomancer-workbench-support/web/What%20is%20georeferencing?>

BioGeomancer workbench is available as either a **single georeference** workbench or a **batch** workbench.

In the single georeference workbench you can type or paste a locality description. If at least one possible location is found, all possible locations that appear to match the locality description are shown on the map. They are also listed in the legend and can be edited.

You can choose the location that is appropriate (by deleting all others); move it around to improve its position; and you can also edit the extent of uncertainty around that point. Finally you can copy the georeference to the clipboard and paste it into another program.<sup>84</sup>

The BioGeomancer research consortium is coordinated by the University of California at Berkeley, funded by the Gordon and Betty Moore Foundation,<sup>85</sup> and includes a worldwide collaboration including Australian Biodiversity Information Services and Australian Museum.<sup>86</sup>

### Function

- Data cleaning and manipulation
  - Georeferencing – applying latitude and longitude
- Visualisation tools
  - Maps
- Non-biological data
- User interface
  - Personal and institutional
  - Raw data and visual presentation

### Why use this tool?

Over the past 250 years, biologists have gone into the field to collect specimens and associated environmental information documenting the range of life. The results of these explorations are an irreplaceable archive of Earth's biological diversity that plays a fundamental role in generating new knowledge and guiding conservation decisions. Yet, roughly one billion specimen records, and even more species observation records, remain practically unusable in their current form.<sup>87</sup>

Georeferenced biocollection data is in high demand. Mapping species occurrence data is fundamental to describing and analysing biotic distributions. This information is also critical for conservation planning, reserve selection, monitoring, and examining the potential effects of climate change on biodiversity.<sup>88</sup>

### Who will use this tool?

BioGeomancer expects to “bring the cost to value ratio down to the point where every collection that seeks to make its data public will also seek to georeference those records”.<sup>89</sup>

- Data creation
- Data capture
  - Curators
- Data providers

---

<sup>84</sup> <http://groups.google.com/group/biogeomancer-workbench-support/web/What%20is%20the%20BioGeomancer%20Workbench>

<sup>85</sup> <http://code.google.com/p/biogeomancer-core/>

<sup>86</sup> <http://www.biogeomancer.org/about.html>

<sup>87</sup> <http://www.biogeomancer.org/about.html>

<sup>88</sup> <http://www.biogeomancer.org/about.html>

<sup>89</sup> <http://www.biogeomancer.org/about.html>

- Institutions
- Private collections
- Casual users
- Data users
  - Expert
  - Interest groups
- ALA infrastructure

#### How will the tool be used?

- Online tool
- User input is required for the single georeferenced mode - location description in English
- Batch submission – under development

#### Where in the data chain could this tool be used?

- Data source
- ALA central
- User's machine

#### When could this tool be used?

- Before data is made available to ALA
- While data is stored with ALA
- At the time of a user request

#### Availability

- BioGeomancer: <http://www.biogeomancer.org/>
- BioGeomancer workbench: <http://bg.berkeley.edu/latest/>
- Help: <http://groups.google.com/group/biogeomancer-workbench-support/web/biogeomancer-workbench-help>
- BioGeomancer Core: <http://code.google.com/p/biogeomancer-core/>
- BioGeomancer Workbench development: <http://code.google.com/p/biogeomancer-workbench/>
- Article in PLoS Biology 4 November 2006: <http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=1637066>
- Cost: free
- Copyright: <http://www.biogeomancer.org/copyright.html>
- Licence: Unknown

#### Comments

The BioGeomancer project has also developed a [Guide to best practices for georeferencing](#)<sup>90</sup>

A single georeferenced request can take 20 seconds or more. These times may improve if hosted locally.

Under development is a Georeferencing Web Service that “accepts a DwC record and returns that record along with all georeferences for it generated by the BioGeomancer Core API”.<sup>91</sup>

<sup>90</sup> <http://www.gbif.org/prog/digit/Georeferencing>

<sup>91</sup> <http://code.google.com/p/biogeomancer-core/wiki/WebServicesRequirements>

## Gazetteers

Summary	
Type of tool	Searchable databases
Function	Georeferencing
Online / Desktop	Online and desktop
Computer infrastructure	
Development status	
Time of use	In preparing data, while checking data, and at time of user request
Licence	

**Gazetteers** are a source of place names giving feature identification and geographic coordinates.

### Description

Gazetteers are available for Australia from a range of sources.

#### Gazetteer of Australia

This is available as a CD. There is also an [Online Place Name Search](#).

The Gazetteer of Australia provides information on the location and spelling of 322,328 geographical names across Australia as of March 2006. Supply of data is coordinated by the [Intergovernmental Committee on Surveying and Mapping \(ICSM\)](#) and derived from state, territory, and Australian government agencies. Copyright of the gazetteer data and postcode data resides with the relevant [state, territory and Australian government agencies](#) within Australia who are custodians of the data.

The gazetteer fields include:

Field	Description
Record ID	Unique feature identifier for each feature.
Authority ID	Custodian state or territory.
State ID	State or territory that feature falls in.
Name	Name of the feature.
Feature Code	Code indicating the type of feature - ( <a href="#">Feature Codes Listing</a> )
Status	Indicates if the name is authorised.
Variant name	Variant or alternative name used for feature.
Postcode	Australia Post Postcode for that feature.
Concise Gazetteer	Indicates if feature is included in the Concise Gazetteer.
Longitude	Longitude of the feature in decimal degrees.
Latitude	Latitude of the feature in decimal degrees.
100K map number	1:100,000 scale map number in which feature is located.
CGDN	Indicates if the place name can be used in the state.au second level domains by community website portals that reflect community interests.

**Please note:** Any organisation or individual wanting to use the Gazetteer data in a similar capacity as the Online Place Name Search or any other online application requires an Internet use licence.<sup>92</sup>

The custodians of the place name data do not guarantee that the data is free from errors and omissions.

### **Product Specifications**<sup>93</sup>

**Coverage:** Australia

**Currency:** 2006

**Coordinates:** Geographical

**Datum:** GDA94

**Format:** Fixed width ASCII and Microsoft Access Database

**Medium:** CD-ROM

**Forward Program:** Annual revision

### Australian National Placenames Survey

The aim of the Australian National Placenames Survey (ANPS) is to prepare a national database of geographical names. It forms a supplement to the National Gazetteer, and formed on established principles within the disciplines of history, geography, linguistics, cultural studies, surveying and mapping. It will record all known Australian names, documenting their pronunciation, generic class, status (gazetted, obsolete, non-gazetted, sensitive, disputed etc.), origin, meaning, history, cultural significance (of both name and site), and map reference and location. The database is accessible for public enquiry via the Web and will be available to enable the production of place name dictionaries, both local and national.<sup>94</sup>

### Maritime Gazetteer of Australia

The Maritime Gazetteer of Australia (MGA) is a searchable database containing all the place names used in Australia's official navigational charts.

Positions relating to a feature name are displayed to the nearest minute of latitude and longitude together with a limited number of other attributes, such as all the charts that include each name. The MGA does not provide the gazetted geographic position of features.

Place names contained in Australia's official navigational charts for the Australian mainland, offshore territories, Antarctica and PNG are in the MGA. The MGA is updated and enhanced as new charts are published.

The MGA database currently displays the following fields:

- Feature Name: The Name as it appears on the official navigational chart.
- Feature Type: This shows the Code of the feature type which the name represents. For details see the [Feature Code Index](#).
- Latitude and Longitude: The position of the name on the chart. *Note this is not necessarily the position of the feature itself, which may only lie partially on the chart (for example an ocean).*
- Charts on which the name appears.
- Grouping: This field shows the association of a collection of features to a group (for example a collection of Islands which form an Island Group).

<sup>92</sup> <http://www.ga.gov.au/nmd/products/thematic/gaz2006.jsp>

<sup>93</sup> <http://www.ga.gov.au/nmd/products/thematic/gaz2006.jsp>

<sup>94</sup> <http://www.anps.org.au/about.html>

- Feature ID - a unique identifier for each name in the database. The MGA can be searched online by Feature Name.<sup>95</sup>

#### Australian Heritage Places Inventory

The Australian Heritage Places Inventory contains summary information about places listed in State, Territory and Commonwealth Heritage Registers. The information is updated regularly. The website includes a search, results, metadata and links to the parent databases.<sup>96</sup>

#### Antarctic Gazetteer

A catalogue of Antarctic Names from the SCAR Composite Gazetteer of Antarctica, the Australian Antarctic Gazetteer and lists of other less formal names.

The **Australian Antarctic Names and Medal Committee** (AANMC) is the authority on Australian names in the Australian Antarctic Territory, and the Territory of Heard Island and the McDonald Islands.<sup>97</sup>

#### States and Territories

There are also gazetteers for the states and territories. See reference to Online Gazetteers below.

### **World-wide Gazetteers**

#### BioGeomancer

BioGeomancer is an online georeferencing tool, with an inbuilt gazetteer, and returns the geographic coordinates of a location submitted as plain text.

#### GEBCO Names of Undersea Features

GEBCO – General Bathymetric Chart of the Oceans. Undersea feature names to appear on sheets/products of GEBCO and Regional International Bathymetric Charts projects, and on international nautical charts are selected by the GEBCO Sub-Committee on Undersea Feature Names (SCUFN). All names selected by SCUFN are contained in the GEBCO Gazetteer of Undersea Feature Names, available as an Excel spreadsheet and as a PDF file.<sup>98</sup>

FuzzyG – The Fuzzy Gazetteer - enables you to find geographic features even when you do not know their exact names. Returns a list of similar names. Datum is WGS84 and precision to nearest minute. Original coordinate system Degrees, Minutes, Seconds (DMS). Based on NIMA GeoNET names server data.<sup>99</sup>

Falling Rain – a gazetteer site that gives latitude and longitude in decimal degrees, and elevation in feet. Derived from NIMA data, this is a useful gazetteer for countries where gazetteers are generally lacking eg. Papua New Guinea.

The Alexandria Digital Library Gazetteer - Gazetteer that's especially useful for historic and variant names. Datum is WGS84. Data based on US NIMA, precision usually to the nearest minute. Original coordinate system= DMS.<sup>100</sup>

<sup>95</sup> <http://www.hydro.gov.au/tools/mga/mga.htm>

<sup>96</sup> <http://www.heritage.gov.au/ahpi/about.html>

<sup>97</sup> [http://data.aad.gov.au/aadc/gaz/search\\_names.cfm](http://data.aad.gov.au/aadc/gaz/search_names.cfm)

<sup>98</sup> <http://www.ngdc.noaa.gov/mgg/gebco/underseafeatures.html>

<sup>99</sup> <http://www.herpnet.org/Gazetteer/GeorefResources.htm>

<sup>100</sup> <http://www.herpnet.org/Gazetteer/GeorefResources.htm>

Maporama - For finding coordinates for street addresses anywhere in the world. Coordinate precision to the nearest minute.<sup>101</sup>

#### **Function**

- Data cleaning and manipulation
  - Data cleaning – spelling, misnaming
  - Data validating – taxonomy, geography
  - Georeferencing – applying latitude and longitude
- Provider interaction
  - Data preparation
- Metadata
  - Creation
- Non-biological data
- User interface
  - Personal and institutional use
  - Raw data and visual presentation

#### **Why use this tool?**

- To enable consistent georeferencing of collections where placenames are known

#### **Who will use this tool?**

- Data creation
- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
  - Casual users
- ALA infrastructure

#### **How will the tool be used?**

- These gazetteers can be online searchable databases, or CDs containing databases for desktop use
- Generally user input is required.
- BioGeomancer will run batch jobs<sup>102</sup>

#### **Where in the data chain could this tool be used?**

- Data source
- ALA central
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- While data is stored with ALA
- At the time of a user request

---

<sup>101</sup> <http://www.herpnet.org/Gazetteer/GeorefResources.htm>

<sup>102</sup> see Chapman, A.D. and J. Wieczorek (eds). 2006. Guide to Best Practices for Georeferencing. Copenhagen: Global Biodiversity Information Facility. <http://www.gbif.org/prog/digit/Georeferencing>

## Availability

### Links

- National Library of Australia - [Gazetteers of the World and Beyond](http://www.nla.gov.au/map/worldgazetteers.html)
- National Library of Australia - [Gazetteers of Australia](http://www.nla.gov.au/map/gazetteers.html)
- ICSM – Intergovernmental Committee on Surveying and Mapping - [Gazetteers](http://www.icsm.gov.au/cgna/websites.html)
- BioGeomancer links to gazetteers  
[http://www.biogeomancer.org/bg\\_library/links/gazetteers/](http://www.biogeomancer.org/bg_library/links/gazetteers/)
- GEBCO Names of Undersea Features  
<http://www.ngdc.noaa.gov/mgg/gebco/underseafeatures.html>

### Online Gazetteers

- Geoscience Australia, Place Name Search  
<http://www.ga.gov.au/map/names/>
- Geoscience Australia, Gazetteer of Australia (CD)  
<http://www.ga.gov.au/nmd/products/thematic/gaz.jsp>
- Australian National Placenames Survey  
<http://www.anps.org.au/index.html>
- Australian Heritage Places Inventory  
<http://www.heritage.gov.au/ahpi/index.html>
- Australian Hydrographic Service, Maritime Gazetteer of Australia  
<http://www.hydro.gov.au/tools/mga/mga.htm>
- Australian Antarctic Gazetteer  
[http://data.aad.gov.au/aadc/gaz/search\\_names.cfm](http://data.aad.gov.au/aadc/gaz/search_names.cfm)
- ACT Land Information Centre Place Names  
[http://apps.actpla.act.gov.au/actlic/places/search/originsSearch\\_new.htm](http://apps.actpla.act.gov.au/actlic/places/search/originsSearch_new.htm)
- Queensland Place Names  
[http://www.nrm.qld.gov.au/property/place\\_names.html](http://www.nrm.qld.gov.au/property/place_names.html)
- New South Wales Geographical Names Register  
<http://www.gnb.nsw.gov.au>
- Victoria Place Names  
<http://services.land.vic.gov.au/vicnames/>
- Aboriginal Placenames of Victoria  
<http://vaclang.ozhosting.com/>
- South Australia Gazetteer  
<http://www.placenames.sa.gov.au>
- Western Australia Geographical Names (database information)  
[http://www.landgate.wa.gov.au/corporate.nsf/web/Western Australian Place Names and Addressing](http://www.landgate.wa.gov.au/corporate.nsf/web/Western%20Australian%20Place%20Names%20and%20Addressing)
- Northern Territory Geographical Names  
<http://www.placenames.nt.gov.au>
- Tasmania Official Place Names (contact information)  
<http://www.dpiw.tas.gov.au/inter.nsf/WebPages/JGAY-53N59M?open>
- BioGeomancer includes a gazetteer  
<http://www.biogeomancer.org/>
- GEBCO – General Bathymetric Chart of the Oceans  
<http://www.ngdc.noaa.gov/mgg/gebco/underseafeatures.html>



- FuzzyG - The Fuzzy Gazetteer  
<http://tomcat-dmaweb1.jrc.it/fuzzyg/query/>
- Falling Rain  
<http://www.fallingrain.com/>
- Alexandria Digital Library Gazetteer  
<http://webclient.alexandria.ucsb.edu/>
- Maporama  
<http://world.maporama.com/>
- Costs: Free for online. Licence, for example, for internet use from Geoscience Australia is up to \$1620.  
<http://www.ga.gov.au/nmd/products/purchasing/licencefees.jsp#gaz>
- Licence and copyright: Generally, the data is copyright of the institution or government holding the data

### Comments

Historical gazetteers have an important role for georeferencing collections where place names may have changed or disappeared.

See also: Perry-Castañeda Library Map Collection, Historical Maps of Australia and the Pacific,  
[http://www.lib.utexas.edu/maps/historical/history\\_austral\\_pacific.html](http://www.lib.utexas.edu/maps/historical/history_austral_pacific.html)

Gazetteers are databases. The ALA may choose to host its own rather than to use the services of online gazetteers.

## GEOLocate

Summary	
Type of tool	Application
Function	Georeferencing
Online / Desktop	Desktop
Computer infrastructure	Windows, Microsoft .NET framework, web browser
Development status	Active. Version 2.13, June 2005
Time of use	Data preparation
Licence	Non-transferable, not to be copied

Georeferencing software for natural history collections.

### Description

**GEOLocate** is a software tool that enables researchers to easily assign geographic coordinates to a descriptive string of locality information, visualize the location, and make corrections as necessary.<sup>103</sup>

Georeferencing begins by standardizing the locality description string into a common terms format. Once standardized the locality string is parsed into key geographic identifiers. Some example geographic identifiers used by GEOLocate include the occurrence of named places, navigable river miles, highway names, water body names, legal locations and displacement patterns. These identifiers within the string are used to determine geographic coordinates from database lookups and geographic calculations. The resulting coordinates are ranked based on the type of information found within the string and plotted on the digital map display for user verification, correction and error determination.

#### Features<sup>104</sup>

- Drag and drop coordinate correction
- Batch georeferencing
- File input via .xml, .csv or delimited .txt
- Polygon error determination
- Multiple coordinate determination
- Supports most countries, including Australia (licence from Geoscience Australia required)
- Overview plotting of input datasets

Depending on the quality of the original locality data, georeferencing results can be improved by prior checks of misspelled, missing, incorrect, and/or ambiguous information within the locality dataset.

GEOLocate is funded by the National Science Foundation and developed by Tulane University's Museum of Natural History designed to facilitate the task of assigning geographic coordinates to the locality data associated with natural history collections.<sup>105</sup>

### Function

- Data cleaning and manipulation

---

<sup>103</sup> <http://www.museum.tulane.edu/geolocate/support/Poster2003.ppt>

<sup>104</sup> <http://www.museum.tulane.edu/geolocate/support/Poster2003.ppt>

<sup>105</sup> <http://www.museum.tulane.edu/geolocate/default.aspx>

- Georeferencing – applying latitude and longitude
- User interface
  - Personal and institutional use
  - Raw data and visual presentation

### Why use this tool?

Traditional methods for georeferencing collection data from text descriptions are tedious and time consuming, typically involving finding the locality on either a hardcopy or digital maps, plotting the locality and determining the coordinates.

GEOLocate considerably reduces the time required to georeference locality information.<sup>106</sup>

### Who will use this tool?

- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- No special skills required

### How will the tool be used?

- Manual input and file input
- File input via .xml, .csv or delimited .txt
- Desktop application
- Requires Windows, Microsoft .NET framework, web browser
- User input required or can run as a batch job.

### Where in the data chain could this tool be used?

- Data source

### When could this tool be used?

- Before data is made available to ALA

### Availability

- GEOLocate: <http://www.museum.tulane.edu/geolocate/>
- Request a download: <http://www.museum.tulane.edu/geolocate/order.aspx>
- Manual: <http://www.museum.tulane.edu/geolocate/support/docs.aspx>
- Support: <http://www.museum.tulane.edu/geolocate/support/>
- Contact: [support@museum.tulane.edu](mailto:support@museum.tulane.edu)
- Version 2.15, June 2005
- Licence: Non-transferable, not to be copied:<sup>107</sup>

### Comments

Requires a user licence from Geoscience Australia for Australian data.

<sup>106</sup> <http://www.museum.tulane.edu/geolocate/support/Poster2003.ppt>

<sup>107</sup> <http://www.museum.tulane.edu/geolocate/order.aspx>

## MaNIS Georeferencing Calculator

Summary	
Type of tool	Online application
Function	Georeferencing
Online / Desktop	Online
Computer infrastructure	Requires Java
Development status	Rev. 25 April 2007
Time of use	When preparing data
Licence	Creative Commons Attribution-Share Alike 3.0 Licence

The **MaNIS Georeferencing Calculator** determines the error associated with given latitude and longitude coordinates and offsets.

### Description

The MaNIS Georeferencing Calculator is a Java applet created as a tool to aid in the georeferencing of descriptive localities such as those found in museum-based natural history collections. It was specifically designed for the Mammal Networked Information System (**MaNIS**) Project and has been adopted as well by **HerpNet** and **ORNIS**. The application makes calculations using the methods described in the **Georeferencing Guidelines**.<sup>108</sup>

The results can be copied onto the system clipboard using standard copy and paste key combinations. There is also a tab-delimited record of the data for the current calculation. These data can be copied all at once and pasted into an Excel spreadsheet or into a record in Access that has the same field order.<sup>109</sup>

This application was originally written by John Wieczorek. Later versions benefited from contributions from Qinghua Guo, Carmen Boureau, and Craig Wieczorek.<sup>110</sup>

### Function

- Data cleaning and manipulation
  - Georeferencing
- Provider interaction
  - Data preparation
- Non-biological data
- User interface
  - Personal use
  - The output is visually presented, and in a form that can be copied directly

### Why use this tool?

- To determine uncertainty associated with a specified location.

### Who will use this tool?

- Data creation
- Data capture
  - Curators

---

<sup>108</sup> <http://manisnet.org/CoordCalcManual.html>

<sup>109</sup> <http://manisnet.org/CoordCalcManual.html>

<sup>110</sup> <http://manisnet.org/gc.html>

- Data providers
  - Institutions
  - Private collections
  - Casual users

#### **How will the tool be used?**

- Needs latitude and longitude. Degrees/minutes/seconds or decimal degrees.
- Online
- No apparent batch mode
- User input required

#### **Where in the data chain could this tool be used?**

- Data source

#### **When could this tool be used?**

- Before data is made available to ALA

#### **Availability**

- MaNIS Georeferencing Calculator: <http://manisnet.org/gci2.html>
- Georeferencing Calculator Manual: <http://manisnet.org/CoordCalcManual.html>
- MaNIS/HerpNet/ORNIS Georeferencing Guidelines: <http://manisnet.org/GeorefGuide.html>
- MaNIS Georeferencing: <http://manisnet.org/search.shtml>
- BioGeomancer: <http://www.biogeomancer.org/>
- Guide to Best Practices for Georeferencing (GBIF publication)  
<http://www.gbif.org/prog/digit/Georeferencing>
- Latest version: Revision 25 April 2007
- Copyright: Regents of the University of California
- Licence: Creative Commons Attribution-Share Alike 3.0 License  
<http://creativecommons.org/licenses/by-sa/3.0/>

#### **Comments**

A good knowledge of the limitations of the data increases the value of that data. This calculator addresses that in an objective manner, and gives a single unqualified value for the error distance.

I am surprised that there is no rounding to the error distance. For example, the error distance in an example that I used was 143.44 km! Why not round to 150 km?

The entering of data is quite labour intensive. If this were to be used for large collections, it could get quite tedious.

## infoXY

Summary	
Type of tool	Application
Function	Reverse georeferencing
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Developed
Time of use	Before data is with ALA , while data is with ALA, and as a post process
Licence	

A simple tool that identifies geographic coordinates.

### Description

This tool was developed by CRIA with the aim of helping biological collections in identifying geographic coordinates.

*infoXY* accepts geographic coordinates and returns information such as the name of the country, state or administrative region. <sup>111</sup>

id	longitude	latitude	country	admin	typeadmin	county (only Brasil)
1	148.6943206	-32.2719298	Australia	New South Wales	State	
2	131.0833282	-25.4	Australia	Northern Territory	Territory	

*Results of a search for two Australian locations* <sup>112</sup>

*infoXY* accepts data as longitude and latitude entered into a web-based text-box. The results area shown on a MapServer map of the world, in an HTML table and as an Excel file.

### Function

- Data cleaning
  - Data validating – geography
- Visualisation tools
  - Maps
- User interface
  - Personal or institutional use
  - Raw data or visual presentation

### Why use this tool?

- To give a location to a latitude and longitude reference.
- This could be used as a means by which the ALA could introduce uncertainty to sensitive information

<sup>111</sup> <http://splink.cria.org.br/infoxy?criaLANG=en>

<sup>112</sup> <http://splink.cria.org.br/infoxy?criaLANG=en>

### Who will use this tool?

- Data capture
  - Curators – specimens
- Data providers
  - Institutions
- Data users
  - General public
- ALA infrastructure
- No special skills required

### How will the tool be used?

- Input data format is comma separated **ID, longitude, latitude**
- Output data is HTML, xls, and displayed on a MapServer map
- Desktop application
- User input required
- Batch job can accept many coordinates at one time
- This tool could be modified to include Australian localities
- *infoXY* could also be used as a “de-sensitive” or dithering tool (see comment below)
- *Data Cleaning* includes/links to this tool<sup>113</sup>

### Where in the data chain could this tool be used?

- Data source
- ALA central
- User’s machine

### When could this tool be used?

- Before data is made available to ALA
- While data is stored with ALA
- As a post process, after data is with the user

### Availability

- infoXY: <http://splink.cria.org.br/infoxy?criaLANG=en>
- CRIA - Centro de Referência em Informação Ambiental: <http://splink.cria.org.br>
- Licence: negotiate with CRIA to modify and use this tool locally

### Comments

For Australia, the results of a search returns the relevant state or territory. This uncertainty may be too vague. infoXY should be modified to include a finer grain locality dataset.

In order to dither the accurate latitude/longitude of sensitive data, a modified form of *infoXY* could be used. By including local government areas, or perhaps state geo-regions<sup>114</sup> (such as Central Tablelands, North Kennedy, etc) in the location database, an accurate geographic coordinate could be de-identified to a less accurate, but relevant geographic area.

---

<sup>113</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

<sup>114</sup> Arthur Chapman, Australian Biodiversity Information Services, January 2008

## Data analysis



## ANHAT – Australian National Heritage Assessment Tool

Summary	
Type of tool	Application
Function	Taxonomic validating and modelling
Online / Desktop	Desktop
Computer infrastructure	Windows, Access, GIS
Development status	Active development
Time of use	At time of user request
Licence	Contact Australian Department of Environment

An analysis and assessment tool for Australian plant and animal species. ANHAT enables the comparative heritage value of different places.

### Description

ANHAT is a tool that assists in the comparative analysis of heritage attributes such as species richness and distribution across bioregions.<sup>115</sup>

ANHAT is designed to perform basic comparative analyses on the presence or absence of taxa across multiple genera, families or orders. It displays the result in a geographic information system (ArcGIS) as a simple shaded map of Australia<sup>116</sup> where each grid cell represents a 10km x 10km area.

This is a specialised tool that has components of interest to the ALA

- Species names management – endorsed names, synonyms, misspellings, truncation
- Species phylogeny using hierarchies downloaded from AFD - Australian Faunal Directory and APNI - Australian Plant Names Index.

ANHAT has been developed by the National Heritage Assessment section of the Australian Department of Environment, Water, heritage and the Arts. They are open to sharing tools with the ALA.

### Function

- Data cleaning and manipulation
  - Data cleaning – spelling, misnaming
  - Data validating – taxonomy, geography
- Analysis tools
  - Simple – distribution
- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Raw data and visual presentation

### Why use this tool?

- Analysis of species distributions

<sup>115</sup> <http://www.environment.gov.au/soe/2006/publications/commentaries/heritage/knowledge.html>

<sup>116</sup> [http://www.anu.edu.au/BoZo/Crisp/C&C%20pdfs/Slatyer\\_2005\\_Frogs.pdf](http://www.anu.edu.au/BoZo/Crisp/C&C%20pdfs/Slatyer_2005_Frogs.pdf)

### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens
- Data providers
  - Institutions
  - Private collections
- Data users
  - Expert
  - Interest groups
  - General public
- ALA infrastructure

### **How will the tool be used?**

- Desktop application
- Microsoft Access, ArcGIS (for mapping)
- ANHAT is a large application that could be broken into separate tools for incorporating into the ALA – eg names management
- User input is required
- Many of the processes can be run as batch jobs.

### **Where in the data chain could this tool be used?**

- Data source
- ALA central
- User's machine
- Pathways between these

### **When could this tool be used?**

- Before data is made available to ALA
- While data is stored with ALA
- At the time of a user request

### **Availability**

- Contact: Dan Rosauer [Dan.Rosauer@environment.gov.au](mailto:Dan.Rosauer@environment.gov.au), Australian Department of Environment, Water, Heritage and the Arts
- Access to this in-house program is subject to negotiation with the Australian Department of Environment, Water, Heritage and the Arts
- Further development will be required to adapt parts of this tool to the ALA

### **Comments**

- Requires development for incorporating into ALA
- The Australian Department of Environment, Water, Heritage and the Arts are open to the concept of the ALA making use of some of their software tools.
- The species lists are incomplete. Some taxonomies are not included :
  - ANHAT doesn't include every plant – currently 13 orders, 80 families
  - Includes all vertebrates
  - Includes chosen invertebrates
- 10km x 10km analysis squares allows for very simple spatial queries
- Are the grid squares arbitrarily chosen, or a known standard?

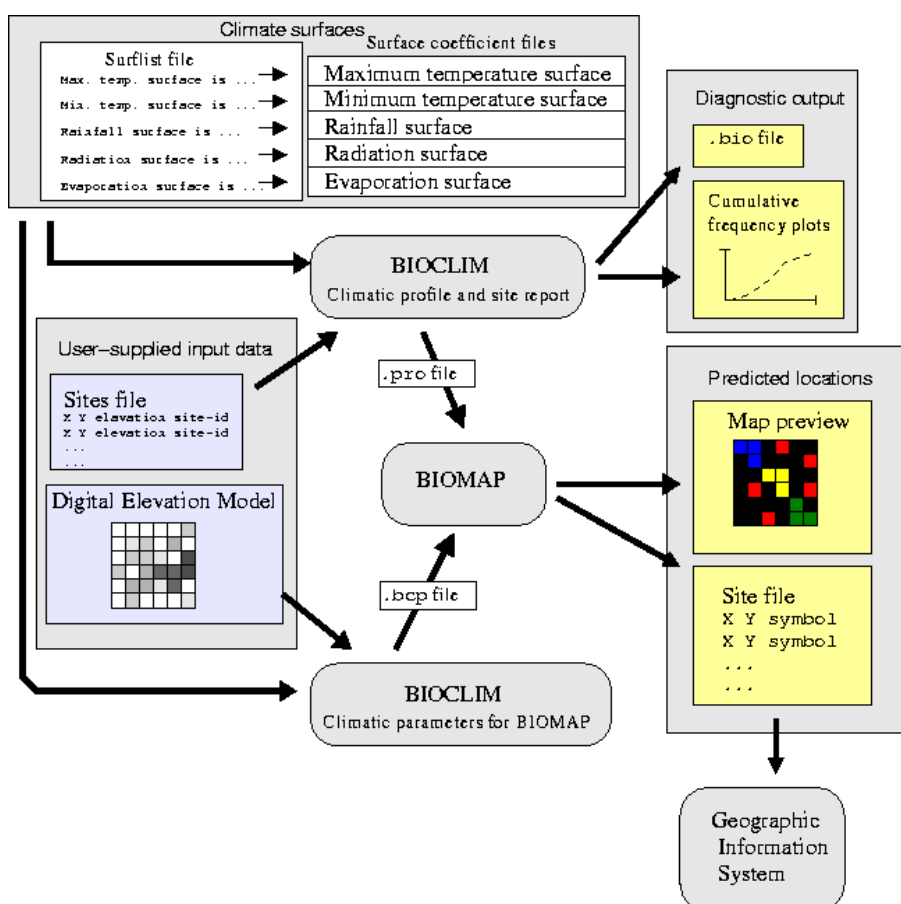
## BIOCLIM – Bioclimatic prediction

Summary	
Type of tool	Application
Function	Species distribution modelling
Online / Desktop	Desktop
Computer infrastructure	Windows, Unix. DEM, GIS may be required.
Development status	Dated. ANUCLIM Version 5.1, December 2000
Time of use	Post process, after data is with user
Licence	

**BIOCLIM** is a bioclimatic prediction system which uses surrogate terms (bioclimatic parameters) derived from mean monthly climate estimates, to approximate energy and water balances at a given location.<sup>117</sup>

### Description

BIOCLIM can produce up to 35 bioclimatic parameters based on the climate variables maximum temperature, minimum temperature, rainfall, solar radiation and pan evaporation. If some of these climate variables are unavailable, fewer bioclimatic parameters are produced.<sup>118</sup>



Using BIOCLIM as a predictive system.<sup>119</sup>

<sup>117</sup> <http://cres.anu.edu.au/outputs/anuclim/doc/bioclim.html>

<sup>118</sup> <http://cres.anu.edu.au/outputs/anuclim/doc/bioclim.html>

<sup>119</sup> <http://cres.anu.edu.au/outputs/anuclim/doc/bioclim.html>

BIOCLIM can either summarise the bioclimatic parameters for a list of sites to produce a species profile, or generate output consisting of the calculated parameters for each point in the input file. These points can be supplied either as a list of coordinates or as a regular grid.<sup>120</sup>

BIOCLIM has a number of advantages over most other models in cases where there is presence-only data (i.e. no absence data), and where the number of known records is extremely low. This makes it a particularly useful modelling system for use with threatened species.<sup>121</sup>

The main drawbacks of BIOCLIM include:<sup>122</sup>

- A tendency to over-predict (i.e. predict larger areas than might otherwise be expected)
- A general restriction to climate-based environmental variables
- An unreliability at fine scales
- The model is sensitive to spatial inaccuracies in the data; a collection site which is in the wrong place can seriously affect the modelled distribution.
- There are many areas which may have climates similar to the species being modelled where the species does not occur. This is called the fundamental niche (Godown and Peterson 2000 ([http://www.specifysoftware.org/Informatics/bios/biostownpeterson/GP\\_BC\\_2000.pdf](http://www.specifysoftware.org/Informatics/bios/biostownpeterson/GP_BC_2000.pdf))). The areas where the species are actually found within the fundamental niche are called the realised niche.
- The fundamental and realised niches are different usually because of factors such as hard boundaries like major rivers and mountain ranges, clearing, land use change etc.
- BIOCLIM tends to overestimate the extent of modelled distributions. However because it is common to only have sampled a small subset of a species true range, models tend to be reasonable guesses at actual distributions.

#### Function

- Analysis tools
  - Simple – distribution
- Non-biological data
  - Environmental data
- User interface
  - Personal use
  - Raw data

#### Why use this tool?

The parameters derived by BIOCLIM (eg Mean Temperature of Warmest Period, Precipitation of Driest Quarter) are useful in determining the climatic envelope for plant and animal species.<sup>123</sup>

#### Who will use this tool?

- Data users
  - Expert
- Special skills required

#### How will the tool be used?

When running BIOCLIM with the species data as input, the input file needs to contain the independent variables used in the creation of the surface coefficient files. For BIOCLIM, this is

---

<sup>120</sup> <http://cres.anu.edu.au/outputs/anuclim/doc/intro.html>

<sup>121</sup> <http://www.biomaps.net.au/biomaps/help.jsp>

<sup>122</sup> <http://www.biomaps.net.au/biomaps/help.jsp>

<sup>123</sup> <http://cres.anu.edu.au/outputs/anuclim/doc/intro.html>

usually just the latitude (or northing), longitude (or easting) and elevation. This information is usually supplied in the form of a plain-text file.

Your input file of site data should contain lines only for known presences of the species. BIOCLIM is not designed to use known absences of a species, nor is it designed to use abundance information that you might have for your sites.<sup>124</sup>

- When using BIOCLIM as a predictive system a digital elevation model (DEM) is required.<sup>125</sup>
- A list of sites is required, each site accurately describing its position and elevation.
- Desktop application
- Windows, Unix (not Linux?)
- GIS may be required
- BIOCLIM is generally hosted inside another program such as ANUCLIM or BioMaps
- User input is required

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- As a post process, after data is with the user

#### **Availability**

- ANUCLIM: <http://cres.anu.edu.au/outputs/anuclim.php>
- ANU Centre for Resource and Environment Studies: <http://cres.anu.edu.au>
- Documentation: <http://cres.anu.edu.au/outputs/anuclim/doc/biocl原因.html>
- ANUCLIM Version 5.1, December 2000<sup>126</sup>
- Support: [anuclim-support@cres.anu.edu.au](mailto:anuclim-support@cres.anu.edu.au)
- Cost: (ANUCLIM) \$1980 <http://cres.anu.edu.au/outputs/orderform-aust.php>

#### **Comments**

- BIOCLIM is a component of ANUCLIM, a software package of programs that enable the user to obtain estimates of monthly mean climate variables, bioclimatic parameters, and indices relating to crop growth. The main components of the package are ESOCIM, BIOCLIM and GROCLIM. They use mathematical descriptions of the way a set of climate variables change across a region in order to estimate those climate variables, or parameters derived from them, at user specified points within the region.<sup>127</sup>
- BIOCLIM was developed in 1986 by Nix and is showing its age.
- There are newer methods that outperform BIOCLIM and other older modelling methods.<sup>128</sup>
- BIOCLIM is a component of BioMaps

---

<sup>124</sup> <http://cres.anu.edu.au/outputs/anuclim/doc/biocl原因.html>

<sup>125</sup> <http://cres.anu.edu.au/outputs/anuclim/doc/biocl原因.html>

<sup>126</sup> <http://cres.anu.edu.au/outputs/anuclim/changes.html>

<sup>127</sup> <http://cres.anu.edu.au/outputs/anuclim/doc/intro.html>

<sup>128</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151, available at: <http://www.blackwell-synergy.com/toc/eco/29/2>

## Biodiverse

Summary	
Type of tool	Application
Function	Biodiversity modelling
Online / Desktop	Desktop and online
Computer infrastructure	Perl based desktop and an online web server
Development status	Experimental
Time of use	As a post process, when data is with user
Licence	Unknown

**Biodiverse** is a powerful spatial analysis tool for analysing multi species patterns of species distribution.

### Description

Analyses are constructed to account for many of the limitations inherent to museum data (eg from herbaria), in that they do not assume that an absence from the data set is a true absence.<sup>129</sup>

Biodiverse is able to use datasets of sample point data for any practical grid size.

Biodiverse can then undertake a quantitative analysis of:

- Species richness
- Endemism
- Neighbourhood analysis
- Compositional dissimilarities
- Sampling adequacy

Biodiverse can also accept Nexus formatted tree data and is able to create clade area relationships.

### Function

- Visualisation tool
  - Maps
- Analysis tool
  - Complex
- User interface
  - Personal use
  - Data is displayed visually through its own GIS interface, or can be saved as a spreadsheet.

### Why use this tool?

- Analysis of species distributions

### Who will use this tool?

- Data users
  - Expert
- Requires special skills

---

<sup>129</sup> <http://www.biodiverse.unsw.edu.au/>

**How will the tool be used?**

- Uses datasets of any species with point data of any chosen grid size
- Acceptable data formats: CSV, ORGD, XML
- An online web-based tool and a stand-alone desktop program which runs in Perl
- User input is required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- Biodiverse: <http://www.biodiverse.unsw.edu.au/>
- Online: <http://www.biodiverse.unsw.edu.au/cgi-bin/biodiverse/biodiverse.pl>
- Help: <http://www.biodiverse.unsw.edu.au/help.htm>
- Licence unknown. Developed at UNSW by Shawn Laffan [Shawn.Laffan@unsw.edu.au](mailto:Shawn.Laffan@unsw.edu.au)
- Cost unknown

**Comments**

This is very powerful software although in an experimental state.

## Biodiversity Statistics and Analysis

Summary	
Type of tool	Collection of tools
Function	Diversity analysis
Online / Desktop	Desktop
Computer infrastructure	Various. Excel add-in, stand-alone. Mostly Windows-based
Development status	Various. Some tools are in active development, some are dated.
Time of use	As a post process
Licence	All programs are freely accessible. Copyright varies.

This site describes a number of freely available statistics and analysis packages relevant to diversity data.

### Description

Software that is listed here is available freely on the Internet, or was developed by Roeland Kindt during the preparation of his PhD thesis. The aim of the thesis was to explore methods of analysing biodiversity data using the management by farmers of tree species diversity in Kenya.<sup>130</sup>

The software listed here is a collection of free software available for the analysis of diversity data.<sup>131</sup>

- PopTools - an Excel add-in that allows more complicated analyses to be conducted within Excel.
- EstimateS - provides species accumulation curves, and various estimators of total diversity (gamma diversity).
- PAST - includes many of the functions which are commonly use in palaeontology and palaeoecology. Many of these functions are common to biodiversity analysis as well.
- EcoSim - an interactive program for null model analysis in community ecology.
- Biodiversity Professional - supports the calculation of various diversity indices, clustering and ordination.
- Krebswin - based on the textbook Ecological Methodology by Charles Krebs.
- Multiple linear regression - computes a multiple linear regression and performs tests of significance of the equation parameters using permutations.
- K-Means - a least-squares partitioning method allowing users to divide a collection of objects into K groups.
- DistPCoA - performs Principal Coordinate Analysis with the option of correcting for negative eigenvalues.
- Polynomial RdaCca - performs two forms of canonical analysis, linear and polynomial redundancy analysis and canonical correspondence analysis.
- R - a general statistics package, supporting an environment similar to the S-Plus Package.
- Vegan - conducts many types of ecological and community analysis.
- mgcv - conducts GAM – Generalised Additive Models.
- Instat - a general statistics Package.
- Reading SSC Add-Ins - an Excel add-in that is targeted to more general Types of Analysis.
- ExactS - calculates the exact average species richness for random accumulation of sites.

<sup>130</sup> <http://www.worldagroforestry.org/newwebsite/sites/rsu/resources/biodiversity/Thesis/index.asp>

<sup>131</sup> <http://www.worldagroforestry.org/newwebsite/sites/rsu/resources/biodiversity/software/index.asp>



- ExactSN - calculates the exact average species richness for random accumulation of individuals using the formula for rarefaction.
- SpeciesRichRef - calculates the average species richness for a fixed number of randomly accumulated sites.
- SpeciesRichClus - calculates the average species richness for sites that are randomly clustered in a fixed number of clusters of fixed size.
- DiversityRef - calculates the averages for various diversity indices for a fixed number of randomly accumulated sites.
- RenyiAccum - calculates average accumulation patterns for the Rényi diversity series.
- RenyiRefs - calculate the average values for selected scales of the Rényi diversity and evenness profiles.
- Metapopulation - calculates the average allelic diversity that remains in an age-structured dioecious metapopulation.

### Function

- Data manipulation
- Analysis tools
- User interface
  - Personal use
  - Raw data and visual presentation

### Why use this tool?

Resources and guidelines useful for statistical analysis of diversity data:<sup>132</sup>

- Averages
- Regression and analysis of variance
- Species richness
- Diversity indices
- Ecological distance
- Clustering
- Ordination

### Who will use this tool?

- Data users
  - Expert
  - Interest groups
- Statistics skills are required

### How will the tool be used?

- Desktop applications
- Some programs stand alone, others are Excel add-ins and supplements to other applications.
- User input is required

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- As a post process, after data is with the user

### Availability

---

<sup>132</sup> <http://www.worldagroforestry.org/newwebsite/sites/rsu/resources/biodiversity/index.asp>

- ICRAF - World Agroforestry Centre:  
<http://www.worldagroforestry.org/newwebsite/sites/rsu/index.asp>
- Biodiversity Analysis:  
<http://www.worldagroforestry.org/newwebsite/sites/rsu/resources/biodiversity/index.asp>
- Software:  
<http://www.worldagroforestry.org/newwebsite/sites/rsu/resources/biodiversity/software/index.asp>
- Roeland Kindt Thesis:  
<http://www.worldagroforestry.org/newwebsite/sites/rsu/resources/biodiversity/Thesis/ThesisComplete.pdf>
- Cost: free
- Copyright. Varies, see individual programs.
- No warranty

### **Comments**

Many of these applications are becoming dated. Listed dates are up to 2002, though many have been upgraded since.

## Biodiversity World

Summary	
Type of tool	Set of tools
Function	Data manipulation
Online / Desktop	Desktop
Computer infrastructure	Works with Triana a tools framework in Java. OS independent
Development status	Probably stagnant
Time of use	Post process when data is with user
Licence	

### Description

The **Biodiversity World** project (**BDWorld**) is creating a GRID-based problem solving environment for collaborative exploration and analysis of global biodiversity patterns. Biodiversity information is stored and indexed using complex dynamic concepts such as species names and geographical areas. The project will provide tools to find information resources, and accurately formulate queries and requests to these resources in the face of conflicting species classifications, geographical terms and descriptor terminology.

The Problem Solving Environment is planned for collaborative exploration and analysis of global biodiversity patterns. Example applications would be modeling species distributions against climate change, conservation prioritising and linking evolutionary changes to past climates. The PSE uses Triana to choreograph custom web services into applications.<sup>133</sup>

Triana is a distributed problem solving environment. It combines workflow with data analysis tools for domains including signal, text and image processing. Triana includes a library of tools and users can integrate their own tools, Web, and Grid Services.<sup>134</sup>

Biodiversity World is a joint collaboration between the Universities of Cardiff, Reading and Southampton, and the Natural History Museum. Biodiversity World has been a three-year e-Science Pilot Project funded by the BBSRC.<sup>135</sup>

### Function

- Analysis tools
  - Complex

### Why use this tool?

- Analysis of global biodiversity patterns

### Who will use this tool?

- Data users
  - Expert

### How will the tool be used?

- BDWorld is a desktop application operating within Triana.
- User input is required

---

<sup>133</sup> <http://www.trianacode.org/collaborations/index.html>

<sup>134</sup> <http://www.trianacode.org/collaborations/index.html>

<sup>135</sup> <http://www.trianacode.org/collaborations/index.html>

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- BDWorld: <http://www.bdworld.org/> (stagnant in Nov 2007)

**Comments**

Biodiversity World appears to be no longer available or supported (November 2007).

What are the data requirements?

See Triana.

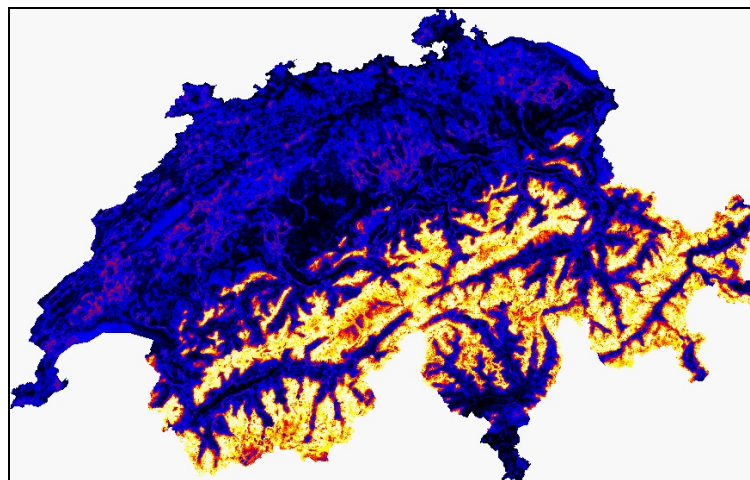
## Biomapper

Summary	
Type of tool	Application
Function	Builds and validates habitat suitability models
Online / Desktop	Desktop
Computer infrastructure	Windows
Development status	Operational version 3.2.2 October 2006
Time of use	Post process data analysis
Licence	Free. Postcard-ware

Using Ecological Niche Factor Analysis, **Biomapper** can build habitat suitability models for any animal or plant.

### Description

Biomapper is a kit of GIS and statistical tools designed to build habitat suitability models and maps for any kind of animal or plant. It is centred on the Ecological Niche Factor Analysis (ENFA) that allows computation of habitat suitability models without the need of absence data.



Habitat Suitability map of *Capra ibex* in Switzerland.<sup>136</sup>

More precisely, it can deal with the following tasks:

- Preparing the ecogeographical maps in order to use them as input for the ENFA (e.g. computing frequency of occurrence map, standardisation, masking, etc.)
- Exploring and comparing them by means of descriptive statistics (distribution analysis, etc.)
- Computing the Ecological Niche Factor Analysis and exploring its output.
- Computing a Habitat Suitability map
- Evaluating it

Biomapper is designed to be autonomous but as it uses the same file format as the GIS software IDRISI, they can transparently work in conjunction.

Biomapper was designed with huge files in mind. It was actually tested with up to 30 maps of 32 Mb each.<sup>137</sup>

<sup>136</sup> <http://www2.unil.ch/biomapper/>

<sup>137</sup> [http://www2.unil.ch/biomapper/what\\_is\\_biomapper.html](http://www2.unil.ch/biomapper/what_is_biomapper.html)

**Function**

- Analysis tools
  - Complex
- Visualisation tools
  - Maps

**Why use this tool?**

- To be able to visually represent habitat suitability

**Who will use this tool?**

- Data users
  - Expert

**How will the tool be used?**

- Requires Windows
- Desktop application
- Field data or database data first needs to be converted to IDRISI-formatted maps
- The habitat suitability maps are validated using independent data
- User input is required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- Biomapper: <http://www2.unil.ch/biomapper/>
- Download: <http://www2.unil.ch/biomapper/products.html>
- Manual: <http://www2.unil.ch/biomapper/Biomapper3-Manual.pdf>
- Version 3.2.2 October 2006
- Licence issues, intellectual property
- Cost: Biomapper is postcard-ware. You can test it freely and, should you find it useful and make use of it for your research, then send a postcard to its author at the following address:  
Dr. Alexandre HIRZEL  
Department of Ecology and Evolution  
University of Lausanne  
CH-1015 LAUSANNE  
SWITZERLAND
- Announcement and discussion lists: <http://www2.unil.ch/biomapper/products.html>
- Restrictions: For Scientific purposes only

**Comments**

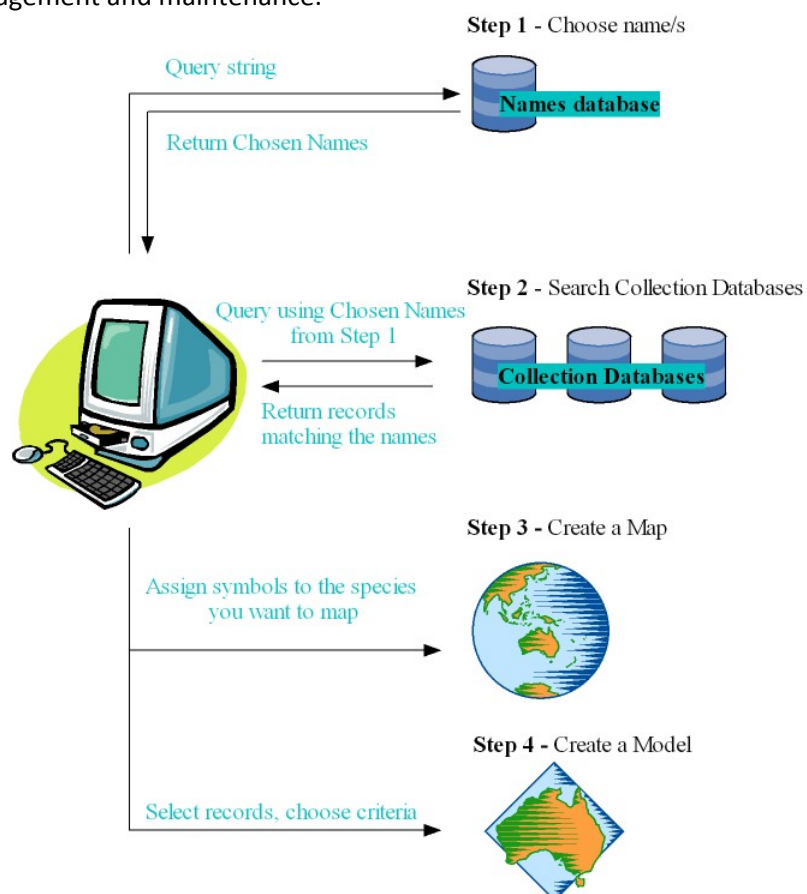
## BioMaps

Summary	
Type of tool	Application
Function	Species mapping, distribution modelling
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Active development
Time of use	At the time of a user request
Licence	

**BioMaps** is a gateway to Australian biodiversity data held by Natural History Institutions. Through this site you can search biodiversity databases and compile maps showing the locations of collected and observed biodiversity specimens.<sup>138</sup>

### Description

BioMaps is a portal to Australian biodiversity data held by government, and scientific organisations in Australia and around the world. It is based on a distributed data structure. This means BioMaps can use the data but the custodial institutions remain responsible for its creation, management and maintenance.<sup>139</sup>



*The four steps to create a BioMaps model<sup>140</sup>*

<sup>138</sup> <http://www.biomaps.net.au/biomaps/>

<sup>139</sup> <http://www.biomaps.net.au/biomaps/about.jsp>

<sup>140</sup> <http://www.biomaps.net.au/biomaps/help.jsp>

BioMaps provides tools for accessing and analysing biodiversity data. It enables visualisation of data, and data analysis products, as a digital map.

BioMaps is being developed as a modular application with components being developed and released over a period of time. The first component is the mapping application which came online in September 2005. Subsequent components will include species mapping and survey planning applications. BioMaps will enable environmental and land planners to access tools for identifying Australian conservation priorities.

Currently BioMaps accesses fauna data from the Online Zoological Collections of Australian Museums (OZCAM) network. This includes the Australian Museum, Museum of Victoria, Museum and Art Gallery of the Northern Territory, South Australian Museum, Commonwealth Scientific and Industrial Research Organisation (CSIRO) Australian National Wildlife Collection, and the Queensland Museum. Other institutions will be made accessible through BioMaps as it develops. By the end of June 2007 it will enable access to flora and fauna data from institutions in Australia and around the world.

The Australian Museum / Rio Tinto Partnership initiated the BioMaps project and provides funding for its development.<sup>141</sup>

#### **Function**

- Visualisation tools
  - Maps
- Analysis tools
  - Simple – distribution
- User interface
  - Personal use
  - Visual presentation

#### **Why use this tool?**

BioMaps will help identify places of biodiversity significance in Australia and be a guide to direct survey and conservation efforts.<sup>142</sup>

#### **Who will use this tool?**

- Data users
  - Expert
  - Interest groups
- Special skills required

#### **How will the tool be used?**

- BioMaps uses either the BIOCLIM or DOMAIN model<sup>143</sup>
- The models can be used with GIS
- Georeferenced species data is available
- Online application
- Browser required
- User input required

---

<sup>141</sup> <http://www.biomaps.net.au/biomaps/about.jsp>

<sup>142</sup> <http://www.biomaps.net.au/biomaps/about.jsp>

<sup>143</sup> <http://www.biomaps.net.au/biomaps/help.jsp>



**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- At the time of a user request

**Availability**

- BioMaps: <http://www.biomaps.net.au/biomaps/>
- Help: <http://www.biomaps.net.au/biomaps/help.jsp>
- Contact: Paul Flemons [Paul.Flemons@austmus.gov.au](mailto:Paul.Flemons@austmus.gov.au)

**Comments**

- Requires accurate scientific name
- Can be slow (1-2 minutes)
- See also review of BIOCLIM and DOMAIN

BioMaps is under active development. A more sophisticated version is currently under test with improved maps and a more functional user interface.

## BRT – Boosted Regression Trees

Summary	
Type of tool	Application
Function	Species modelling
Online / Desktop	Desktop
Computer infrastructure	Windows, Unix, R
Development status	Version 1.6 August 2007
Time of use	As a post process, after data is with the use
Licence	

**BRT** – Boosted Regression Trees, also called stochastic gradient boosting.

### Description

Boosted regression trees combine two algorithms: “boosting” is a method for developing multiple models and combining them; “regression trees” are single models that partition the predictor space into disjoint regions and predict a separate constant value in each of them.

Boosting is used to overcome the inaccuracies of a single model, and makes it possible to model a complex response surface. Regression trees can use continuous and categorical predictor variables, allow for missing data, are not sensitive to outliers, tend to exclude irrelevant variables, and model interactions.

BRT are described in different ways in different disciplines. The foremost interpretation from the machine learning community is that it is a method for finding many rough rules of thumb (i.e. many regression trees) that, when combined, are more accurate than any single rule. The boosting algorithm calls the regression tree algorithm repeatedly, each time giving it a re-weighted version of the data that emphasizes the records that were misclassified in the last round. Finally the suite of trees is combined by weighted averaging. Statisticians have reinterpreted it as a method for developing a regression model in a forward stage-wise fashion, adding small modifications across the model space (via trees) to fit the data better. The final model has numerous terms, each term being a regression tree.

As boosting proceeds, the model complexity increases until eventually it over-fits the data. The number of trees in the boosted model is a natural measure of complexity, and is chosen by measuring prediction accuracy on independent data. This identifies the most complex model that still predicts well, and is based on the trade-off between training error and generalization error.<sup>144</sup>

### Function

- Analysis tools
- User interface
  - Personal use
  - Raw data

### Why use this tool?

---

<sup>144</sup> Friedman and Meulman 2003 in Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151.  
available at: <http://www.blackwell-synergy.com/toc/eco/29/2>

- Species modelling

#### **Who will use this tool?**

- Data users
  - Expert
- Special skills are required

#### **How will the tool be used?**

- BRT is part of the GBM library of R
- Desktop application
- User input required

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- As a post process, after data is with the user

#### **Availability**

R Project for Statistical Computing

- R: <http://www.r-project.org/>
- GBM package (Generalised Boosted Regression Models) Version 1.6-3 (Aug 2007):  
<http://cran.r-project.org/src/contrib/Descriptions/gbm.html>
- GBM manual: <http://cran.r-project.org/doc/packages/gbm.pdf>
- GBM guide: <http://cran.r-project.org/doc/vignettes/gbm/gbm.pdf>
- Licence: GPL-2, GNU Public Licence
- Cost: free

#### **Comments**

For species' distribution modelling, BRT has been compared to other methods.<sup>145</sup>

---

<sup>145</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151. available at: <http://www.blackwell-synergy.com/toc/eco/29/2>

## BRUTO – Generalised Additive Modelling

Summary	
Type of tool	Application
Function	Species modelling
Online / Desktop	Desktop
Computer infrastructure	Windows, Unix, R and S-PLUS
Development status	Commercial and recent freeware
Time of use	As a post process, after data is with the use
Licence	

**BRUTO** is a fast implementation of a Generalised Additive Modelling tool.<sup>146</sup>

### Description

BRUTO fits a generalized additive model (GAM) using an adaptive back-fitting procedure with smoothing splines.

GAMs are multiple regression models in which non-parametric smooth functions are used to model non-linear relationships. They are able to deal with categorical data; can include a mixture of linear and non-linear fitted functions; can model a variety of response types, including binomial and Poisson. A range of alternative smoothers are available.

In addition to identifying which variables to include in the final model, BRUTO identifies the optimal degree of smoothing for each variable. BRUTO also allows specification of a penalty parameter that is applied to the addition of extra variables in the model.

The model selection is based on an approximation to the generalized cross-validation (GCV) criterion, which is used at each step of the back-fitting procedure. Once the selection process stops, the model is backfit using the chosen amount of smoothing.<sup>147</sup>

### Function

- Analysis tools
- User interface
  - Personal use
  - Raw data

### Why use this tool?

- In large data sets it is about 100 times faster at fitting a model than a GAM<sup>148</sup>

### Who will use this tool?

- Data users
  - Expert
- Special skills are required

---

<sup>146</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151. Online appendix to this paper: E4596 <http://www.oikos.ekol.lu.se/app.html>

<sup>147</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151. Online appendix to this paper: E4596 <http://www.oikos.ekol.lu.se/app.html>

<sup>148</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151. Online appendix to this paper: E4596 <http://www.oikos.ekol.lu.se/app.html>

**How will the tool be used?**

- BRUTO is part of the MDA library of R and S-PLUS
- Desktop application
- User input required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

R Project for Statistical Computing

- R: <http://www.r-project.org/>
- MDA package (mixture and flexible discriminate analysis) Version 0.3-2 (Oct 2006): <http://cran.r-project.org/src/contrib/Descriptions/mda.html>
- Manual: <http://cran.r-project.org/doc/packages/mda.pdf>
- Licence: GPL-2, GNU Public Licence
- Cost: free

S-PLUS commercial statistics

- Insightful Corporation: <http://www.insightful.com/>
- S-PLUS Version 8, 2007: <http://www.insightful.com/products/splus/>
- MDA package (mixture and flexible discriminate analysis) Version 0.3-2 (Oct 2006): <http://cran.r-project.org/src/contrib/Descriptions/mda.html>
- Manual: <http://cran.r-project.org/doc/packages/mda.pdf>
- Licence: Contact Insightful Corporation

**Comments**

For species' distribution modelling, BRUTO has been compared to other methods.<sup>149</sup>

---

<sup>149</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151. available at: <http://www.blackwell-synergy.com/toc/eco/29/2>

## CLUZ

Summary	
Type of tool	Application
Function	Reserve planning tool
Online / Desktop	Desktop
Computer infrastructure	Windows, ArcView 3.x, MARXAN
Development status	Active development. Version 1.7, December 2007
Time of use	Post process
Licence	Public domain

**CLUZ** is an ArcView GIS interface that allows users to design protected area networks and conservation landscapes. It can be used for on-screen planning and also acts as a link for MARXAN conservation planning software.<sup>150</sup>

### Description

Most protected area (PA) systems have been influenced by political, economic and aesthetic factors and under-represent many elements of biodiversity. Moreover, most PAs are too small to adequately conserve populations of wide-ranging species and important ecological processes. At the same time, many people partly rely on natural resources for their survival and so the unplanned creation or expansion of PAs can cause problems.<sup>151</sup>

Therefore, there is a general need to develop conservation landscapes that allow the maintenance of biodiversity whilst minimising impacts on the livelihoods of local people (Driver et al, 2003). The best way of achieving this is to use systematic conservation planning techniques, as these allow stakeholders to develop land-use plans in a transparent manner by going through the following steps:<sup>152</sup>

1. Compile data on the biodiversity of the planning region
2. Identify conservation goals for the planning region
3. Review existing conservation areas
4. Select additional conservation areas
5. Implement conservation actions
6. Maintain the required values of conservation areas

CLUZ is an ArcView GIS extension that consists of the CLUZ menu, which contains 25 items, as well as 6 buttons and 2 tools. Each of these elements carries out a different function.<sup>153</sup>

CLUZ is currently being developed at DICE at the University of Kent and is funded by the British Government through their Darwin Initiative for the Survival of Species.<sup>154</sup>

### Function

- Analysis tools
  - Complex
- User interface
  - Personal

<sup>150</sup> <http://www.mosaic-conservation.org/cluz/index.html>

<sup>151</sup> <http://www.mosaic-conservation.org/cluz/overview.html>

<sup>152</sup> <http://www.mosaic-conservation.org/cluz/overview.html>

<sup>153</sup> <http://www.mosaic-conservation.org/cluz/functions.html>

<sup>154</sup> <http://www.mosaic-conservation.org/cluz/index.html>

- Visual presentation

### Why use this tool?

- For reserve planning.
- To be able to present MARXAN data in a form directly suitable for GIS

### Who will use this tool?

- Data users
  - Expert

### How will the tool be used?

#### What is the difference between CLUZ and MARXAN?

CLUZ consists of two parts. The first part acts as an ArcView GIS interface for the MARXAN conservation planning software and the second allows on-screen conservation planning. The interface for MARXAN is designed to be easily understood by beginners and does not give all of the possible MARXAN options. Specifically, it uses default values for some of the cost functions, it assumes that all of the conservation feature targets should be fully met and it does not allow the use of more specialist planning algorithms.<sup>155</sup>

People who are interested in using the more specialised functionality of MARXAN can still use CLUZ to import, display and export the relevant data but should run MARXAN independently, using its Input File Editor software to set the more specialised parameters.<sup>156</sup>

- Desktop application
- User input is required

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- As a post process, after data is with the user

### Availability

CLUZ is an ArcView GIS extension that partly acts as an interface for the MARXAN conservation planning software. This means that you will need to have all three items installed on your computer before you can run CLUZ.<sup>157</sup>

- CLUZ: <http://www.mosaic-conservation.org/cluz/>
- Download CLUZ: <http://www.mosaic-conservation.org/cluz/download.html>
- MARXAN: <http://www.ecology.uq.edu.au/index.html?page=27710>
- ArcView GIS available from ESRI: <http://www.esri.com/>
- CLUZ tutorial: <http://www.mosaic-conservation.org/cluz/tutorial.html>
- FAQ: <http://www.mosaic-conservation.org/cluz/faq.html>
- Version 1.7, December 2007
- Cost: Free. Public domain. May be freely used and redistributed
- No technical support.

### Comments

---

<sup>155</sup> <http://www.mosaic-conservation.org/cluz/faq.html>

<sup>156</sup> <http://www.mosaic-conservation.org/cluz/faq.html>

<sup>157</sup> <http://www.mosaic-conservation.org/cluz/download.html>

## DesktopGarp

Summary	
Type of tool	Framework with tools
Function	Data analysis: environmental niche modelling
Online / Desktop	Desktop
Computer infrastructure	Windows/Intel only
Development status	Last update June 2005
Time of use	Post-process
Licence	Unspecified

**DesktopGarp** is a software package for biodiversity and ecologic research that allows the user to predict and analyse wild species distributions.<sup>158</sup>

### Description

The acronym GARP stands for Genetic Algorithm for Rule-set Production. GARP was originally developed by David Stockwell, at ERIN in the Australian Department of the Environment and enhanced at the San Diego Supercomputer Center.

GARP is a genetic algorithm that creates an ecological niche model for a species that represents the environmental conditions where that species would be able to maintain populations. GARP uses as input a set of point localities where the species is known to occur and a set of geographic layers representing the environmental parameters that might limit the species' ability to survive.<sup>159</sup>

### Function

GARP tries, interactively, to find non-random correlations between the presence and absence of the species and the values of the environmental parameters, using several types of rules. Each rule type implements a different method for building species prediction models. Currently there are four types of rules implemented: atomic, logistic regression, bioclimatic envelope and negated bioclimatic envelope rules.<sup>160</sup>

### Why use this tool?

To predict and analyse the distribution of species.

### Who will use this tool?

- Data users

### How will the tool be used?

- DesktopGarp runs only on Intel/Windows
- Requires Microsoft XML Parser

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

---

<sup>158</sup> <http://www.nhm.ku.edu/desktopgarp/index.html>

<sup>159</sup> <http://www.nhm.ku.edu/desktopgarp/FAQ.html>

<sup>160</sup> <http://www.nhm.ku.edu/desktopgarp/FAQ.html>



- As a post process, after data is with the user.

#### Availability

- Free to download: <http://www.nhm.ku.edu/desktopgarp/index.html>
- Licence not mentioned
- User's Manual: <http://www.nhm.ku.edu/desktopgarp/UsersManual.html>
- Latest version 1.1.6 June 2005

#### Comments

DesktopGarp comes with DatasetManager, a tool to prepare the data for use with DesktopGarp.

##### **What does the DatasetManager application do?**

DatasetManager is a tool that helps the user prepare sets of geographic layers, or datasets, which will be used by DesktopGarp to build species prediction models. It does not do the entire job, though. It just converts the datasets and creates some metadata that is used by DesktopGarp during optimization. It requires support from a Geographic Information System (GIS), such as ESRI's ArcGIS, Arc/Info or ArcView GIS, to create new GARP datasets.<sup>161</sup>

If the DatasetManager can be used with ALA data, then this could possibly make DesktopGarp a powerful tool.

- Should DesktopGarp be part of the ALA toolkit? Or, is this a third party tool?

---

<sup>161</sup> <http://www.nhm.ku.edu/desktopgarp/FAQ.html>

## DOMAIN

Summary	
Type of tool	Application
Function	Species distribution modelling
Online / Desktop	Desktop
Computer infrastructure	DIVA-GIS Version 5
Development status	Dated
Time of use	As a post process, after data is with the user
Licence	

**DOMAIN** is a model that assesses new sites in terms of their environmental similarity to sites of known presence.<sup>162</sup> It has the ability to operate effectively using only presence records and a limited number of biophysical attributes.<sup>163</sup>

### Description

DOMAIN uses a similarity measure by transforming the known occurrences into an environmental space and computing the minimum distance in environmental space from any cell to a known presence of the species. The result of this is a surface which is ranked for likelihood of having a species occur.<sup>164</sup>

At its very simplest, this model generates maps of similarity or distance. For example to predict the potential distribution of a particular taxon, DOMAIN maps those regions which are most similar to areas where the taxon is known to occur. The measure of similarity used in DOMAIN is based on the Gower metric. For any location in the mapping area, the values in the layer files define an environmental coordinate. For example if 3 layers are open containing rainfall, vegetation type and elevation, the environmental coordinate for any location is the set of three cell values from the three layers at that point. The Gower metric defines a means of computing the distance between any two such environmental coordinates.

The application of DOMAIN is not limited to mapping potential distribution of taxa. The problem of similarity mapping arises in other areas. For example it may be useful to map the regions which are least similar to this set of survey sites when selecting new site locations or examining the adequacy of a sampling strategy.

The DOMAIN algorithm is described in Carpenter, G., Gillison, A.N. and Winter, J. (1993). DOMAIN: a flexible modelling procedure for mapping potential distributions of plants and animals, *Biodiversity and Conservation*. **2**, 667-680.<sup>165</sup>

### Function

- Analysis tools
  - Simple – distribution
- User interface

<sup>162</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151, available at: <http://www.blackwell-synergy.com/toc/eco/29/2>

<sup>163</sup> <http://www.springerlink.com/content/pk788x7lw2u76618/>

<sup>164</sup> <http://www.biomaps.net.au/biomaps/help.jsp>

<sup>165</sup> <http://www.biomaps.net.au/biomaps/help.jsp>

- Personal or institutional use
- Raw data

#### Why use this tool?

- For survey design, reserve selection and potential mapping of rare and common species.<sup>166</sup>

#### Who will use this tool?

- Data users
  - Expert
- Special skills required

#### How will the tool be used?

- DOMAIN algorithm is used as a species prediction tool within DIVA-GIS Version 5<sup>167</sup>
- DOMAIN is used within BioMaps<sup>168</sup>
- Desktop application
- User input required

#### Where in the data chain could this tool be used?

- User's machine

#### When could this tool be used?

- As a post process, after data is with the user

#### Availability

DOMAIN is available through DIVA-GIS version 5

- DIVA-GIS: <http://www.diva-gis.org/>
- Command line version AVID-GIS: <http://www.diva-gis.org/avid.htm>
- Download: <http://www.diva-gis.org/down.asp> registration is voluntary
- Documentation: <http://www.diva-gis.org/Materials.htm>
- Manual: [http://www.diva-gis.org/docs/DIVA-GIS5\\_manual.pdf](http://www.diva-gis.org/docs/DIVA-GIS5_manual.pdf)
- 
- Cost: free
- Licence: GNU General Public License (GPL)

#### Comments

Comparisons of several modelling methods including DOMAIN are available:

Tsoar, Asaf; Allouche, Omri; Steinitz, Ofer; Rotem, Dotan; Kadmon, Ronen (2007) A comparative evaluation of presence-only methods for modelling species distribution *Diversity & Distributions* **13**, 397-405(9).<sup>169</sup>  
available at: <http://www.blackwell-synergy.com/toc/ddi/13/4>

Jane Elith, Catherine H. Graham, Robert P. Anderson, Miroslav Dudi'k, Simon Ferrier, Antoine Guisan, Robert J. Hijmans, Falk Huettmann, John R. Leathwick, Anthony Lehmann, Jin Li, Lucia G. Lohmann, Bette A. Loiselle, Glenn Manion, Craig Moritz, Miguel Nakamura, Yoshinori Nakazawa, Jacob McC. Overton, A. Townsend Peterson, Steven J. Phillips, Karen Richardson, Ricardo Scachetti-Pereira, Robert E. Schapire, Jorge Sobero'n, Stephen Williams,

<sup>166</sup> <http://www.springerlink.com/content/pk788x7lw2u76618/>

<sup>167</sup> DIVA-GIS manual [http://www.diva-gis.org/docs/DIVA-GIS5\\_manual.pdf](http://www.diva-gis.org/docs/DIVA-GIS5_manual.pdf)

<sup>168</sup> <http://www.biomaps.net.au/biomaps/>

<sup>169</sup> <http://www.ingentaconnect.com/content/bsc/ddi/2007/00000013/00000004/art00005>

Mary S. Wisz and Niklaus E. Zimmermann (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* **29**: 129-151.  
available at: <http://www.blackwell-synergy.com/toc/eco/29/2>

Pilar A Hernandez, Catherine H Graham, Lawrence L Master, Deborah L Albert (2006) The effect of sample size and species characteristics on performance of different species distribution modeling methods *Ecography*, **29**, No. 5. (October 2006), 773-785. available at: <http://www.blackwell-synergy.com/toc/eco/29/5>

## Eco-Tools

Summary	
Type of tool	Suite of tools
Function	Calculations in ecology and conservation biology
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Last update March 2006
Time of use	As a post process, after data is with the user
Licence	

**Eco-Tools** is a web-accessible means of performing commonly-needed and useful calculations in ecology and conservation biology.<sup>170</sup>

### Description

No special software is required, and all algorithms are open-source and based on standard references.

#### The Eco-Tools

##### Life Tables – Life Table Calculations

This Eco-Tool uses matrix algebra to calculate a variety of statistics based on a table of *survivorship* and *maternity* values by age class.<sup>171</sup>

##### Population Timeseries – Autocorrelograms and Power Spectra

Analytical population Dynamics. Under development.<sup>172</sup>

##### Population Genetics – F-statistics

Calculates the statistical outcome over several generations of the action of evolutionary forces on allelic and genotypic frequencies. The four evolutionary forces act on genetic variation by changing the frequency of alleles in a population, and the distribution of alleles among populations.<sup>173</sup>

##### Population Viability Analysis (PVA) – Simple Count-Based PVA

This Eco-Tool performs a simple count-based population viability analysis. In short, this is the simplest possible form of PVA, based only on a timeseries of censuses of the population in question.<sup>174</sup>

##### Species Diversity – Estimating Species Richness

The Species Richness eco-tool implements a variety of analyses that take as their input a “species by sample” matrix of observed abundances. The analyses include various summary statistics, as well as a number of estimators of true species richness, including Chao 1, Chao 2, ACE, ICE, first-order jackknife, second-order jackknife, and bootstrap. Variances (and hence confidence intervals) for these estimators are calculated based on analytical results (for Chao 1 and Chao 2) or on multiple random resamples with replacement (for everything else). Finally,

<sup>170</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/index.html>

<sup>171</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/Leslie-1-1/Introduction.html>

<sup>172</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/Autocorrelation-1-1/Introduction.html>

<sup>173</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/Fstats-1-1/Introduction.html>

<sup>174</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/PVA1-1-1/Introduction.html>

the eco-tool calculates a sample-based rarefaction curve, with confidence intervals, as well as the individual-based rarefaction curve for comparison. These curves are plotted, and both the figures and the curve data are provided for download.<sup>175</sup>

### **Ordination – Canonical Methods: RDA and CCA**

This Eco-Tool implements two forms of canonical ordination: Redundancy Analysis (RDA) and Canonical Correspondence Analysis (CCA). The main input consists of a matrix where rows represent sites and columns contain either dependent variables (typically observations of species abundances) or candidate predictor variables (measurements of some environmental quality at each site). A second input is the number of predictor variables; this is used to divide the input matrix into its two components. A third input is the number of randomizations used to test the significance of the dependent variables in determining the observations.<sup>176</sup>

### **Taxonomic and Phylogenetic Autocorrelation – Moran's I**

Moran's I autocorrelation statistics based on taxonomy. Under development.<sup>177</sup>

These tools have been developed by Gareth Russell, Columbia University, and are hosted by the New Jersey Institute of Technology.<sup>178</sup> Eco-Tools is open to input from others.

### **Function**

- Analysis tools
- User interface
  - Personal use
  - Raw data and visual presentation

### **Why use this tool?**

This is a set of numerical tools for ecological analysis that are available, free, to anyone with web access. What makes these tools different from other third-party offerings is that they are programmed in *Mathematica* (Wolfram Research), which is a high-level coding language, and the code is available for anyone to inspect. The idea is that the tools can be easily peer-reviewed, corrected, and improved over time.<sup>179</sup>

### **Who will use this tool?**

- Data users
  - Expert
  - Interest groups
- Special skills required

### **How will the tool be used?**

- Data formats vary depending on the tool in use<sup>180</sup>
- Online tools, programmed in Mathematica
- User input required

### **Where in the data chain could this tool be used?**

- User's machine

---

<sup>175</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/RichnessEstimators-2-1/Introduction.html>

<sup>176</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/Ordination/Canonical.html>

<sup>177</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/MoransI/Introduction.html>

<sup>178</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/SiteInfo/sponsors.html>

<sup>179</sup> [http://eco-tools.njit.edu/webMathematica/EcoTools/SiteInfo/history\\_and\\_purpose.html](http://eco-tools.njit.edu/webMathematica/EcoTools/SiteInfo/history_and_purpose.html)

<sup>180</sup> <http://eco-tools.njit.edu/webMathematica/EcoTools/>

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- Eco-Tools: <http://eco-tools.njit.edu/webMathematica/EcoTools/>
- Contact: [admin@eco-tools.net](mailto:admin@eco-tools.net)
- Cost: free

**Comments**

## GBIF MAPA

Summary	
Type of tool	Portal with tools
Function	Species modelling
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Operational.
Time of use	At the time of a user request
Licence	

**GBIF MAPA** is the Global Biodiversity Information Facility Mapping and Analysis Portal Application.

### Description

MAPA consists of a website which allows users to query the GBIF cache using names obtained through the Catalogue of Life and to map and analyse the resultant record set, and Web services that implement the **SGA** and **SRA** analyses.<sup>181</sup>

The analysis options available through MAPA are:<sup>182</sup>

- **Survey Gap Analysis (SGA)** - This tool helps you design a biodiversity survey that will best complement the existing survey effort by identifying those areas least well surveyed in terms of environmental conditions.
- **Species Richness Assessment (SRA)** - Use this tool to provide an estimate, from GBIF data, of the number of species in an area; and to gain insight into the adequacy of sampling based on abundance distributions for those species.
- **Environment Values Extraction (EVE)** - Query a range of environmental layers (eg climate) using GBIF species record point data to create a table showing the environmental values at those points. This data can then be used in your own statistical analyses.

MAPA has been developed through a collaboration between the Australian Museum (AM), University of Colorado (UC), and the New South Wales Department of Environment and Conservation (DEC). Principal Investigators for the project are Paul Flemons (AM) and Rob Guralnick (UC).<sup>183</sup>

### Function

- Visualisation tools
  - Maps
- Analysis tools
  - Complex
- Non-biological data
  - Environmental data
- User interface
  - Personal and institutional use
  - Visual presentation

### Why use this tool?

---

<sup>181</sup> <http://gbifmapa.austmus.gov.au/mapa/about.jsp>

<sup>182</sup> <http://gbifmapa.austmus.gov.au/mapa/about.jsp>

<sup>183</sup> <http://gbifmapa.austmus.gov.au/mapa/about.jsp>



- To be able to analyse and map biodiversity factors through the one portal

Biodiversity surveys are expensive undertakings requiring careful planning, and specialized resources in terms of personnel skills and equipment. It is essential then that the data obtained through new surveys complements existing data and maximizes the usefulness of the new data for conservation planning purposes. Many, if not all, museum collections are characterized by biased sampling resulting from either ad-hoc collection techniques or from planning that is based on ease of access and which only considers geographic, rather than environmental, coverage when locating survey sites. Survey Gap Analysis can be instrumental in reducing bias and thereby more effectively answering the question "If one is interested in obtaining an overall knowledge of the biodiversity (or of a taxon) of an area, and if there are insufficient data, then where should survey data be gathered?" Funk *et al* 2005<sup>184</sup> <sup>185</sup>.

#### Who will use this tool?

- Data users
  - Expert
  - Interest groups
- Special skills required

#### How will the tool be used?

- Online application
- User input required
- GBIF MAPA uses large numbers of specimen records for its modelling and analysis

#### Where in the data chain could this tool be used?

- User's machine

#### When could this tool be used?

- At the time of a user request

#### Availability

- GBIF MAPA: <http://gbifmapa.austmus.gov.au/mapa/>
- Help and FAQ: <http://gbifmapa.austmus.gov.au/mapa/help.jsp>
- Contact: Paul Flemons [paul.flemons@austmus.gov.au](mailto:paul.flemons@austmus.gov.au)

#### Comments

Maps can take up to several minutes to load.

See: Paul Flemons, Robert Guralnick, Jonathan Krieger, Ajay Ranipeta and David Neufeld, A web-based GIS tool for exploring the world's biodiversity: The Global Biodiversity Information Facility Mapping and Analysis Portal Application (GBIF-MAPA), [Ecological Informatics](#) Volume 2, Issue 1, 1 January 2007, Pages 49-60.

---

<sup>184</sup> Funk, V.A, Richardson, K.S., and Ferrier, S 2005. Survey-gap analysis in expeditionary research: where do we go from here? *Biological Journal of the Linnaen Society*, 2005, **85**, 549-567.

available at: [http://www.mnh.si.edu/biodiversity/bdg/Funk\\_Richardson\\_Ferrier.pdf](http://www.mnh.si.edu/biodiversity/bdg/Funk_Richardson_Ferrier.pdf)

<sup>185</sup> <http://gbifmapa.austmus.gov.au/mapa/help.jsp>

## GDM – Generalised Dissimilarity Modelling

Summary	
Type of tool	Set of functions
Function	Modelling
Online / Desktop	Desktop
Computer infrastructure	Windows, R statistical software
Development status	Active
Time of use	Post process, after data is with user
Licence	

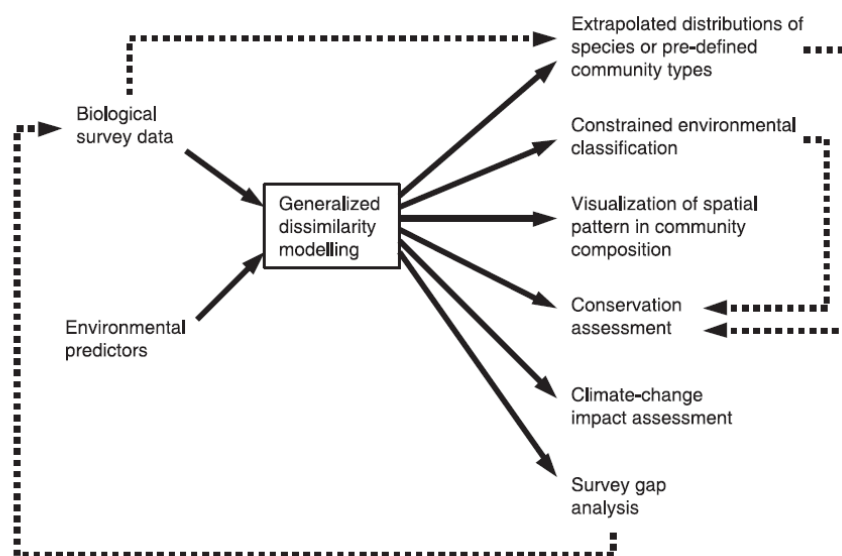
A set of R functions written to perform generalized dissimilarity modelling (**GDM**).<sup>186</sup>

### Description

Generalized dissimilarity modelling (GDM) is a statistical technique for analysing and predicting spatial patterns of turnover in community composition (beta diversity) across large regions.

The approach is an extension of matrix regression, designed specifically to accommodate two types of nonlinearity commonly encountered in large-scaled ecological data sets: (1) the curvilinear relationship between increasing ecological distance, and observed compositional dissimilarity, between sites; and (2) the variation in the rate of compositional turnover at different positions along environmental gradients.

GDM can be further adapted to accommodate special types of biological and environmental data including, for example, information on phylogenetic relationships between species and information on barriers to dispersal between geographical locations.<sup>187</sup>



*Applications of generalised dissimilarity modelling.*<sup>188</sup>

<sup>186</sup> Ferrier and Manion (2007) GDM tutorial

<sup>187</sup> Ferrier, S., Manion, G., Elith, J. and Richardson, K. (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* **13**: 252-264. available at: <http://www.blackwell-synergy.com/toc/ddi/13/3>

<sup>188</sup> Ferrier, S., Manion, G., Elith, J. and Richardson, K. (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* **13**: 252-264. available at: <http://www.blackwell-synergy.com/toc/ddi/13/3>

The GDM tool is a set of functions for performing GDM within the R statistical software environment.<sup>189</sup>

The functions are as follows:<sup>190</sup>

1. `gdm.fit` – used to fit a generalized dissimilarity model to data provided in a site-by-species table and a corresponding table of environmental predictors for the same set of sites.
2. `gdm.summary` – used to produce a summary of the `gdm` model object returned by `gdm.fit`.
3. `gdm.plot` – used to plot the functions fitted to each predictor in a `gdm` model object returned by `gdm.fit`, and to produce scatter-plots of observed versus predicted dissimilarities between site-pairs.
4. `gdm.predict` – uses a `gdm` model object returned by `gdm.fit` to predict compositional (biological) dissimilarity between pairs of sites, based on the values of environmental predictors at these sites.
5. `gdm.transform` – uses a `gdm` model object returned by `gdm.fit` to transform environmental data for any specified set of sites.

Developed by Simon Ferrier and Glenn Manion at NSW Department of Conservation.

#### Function

- Analysis tool
- User interface
  - Personal use
  - Raw data, a GIS version for visual representation is under development

#### Why use this tool?

GDM can be applied to a wide range of assessment activities including visualization of spatial patterns in community composition, constrained environmental classification, distributional modelling of species or community types, survey gap analysis, conservation assessment, and climate-change impact assessment.<sup>191</sup>

#### Who will use this tool?

- Data users
  - Expert
- Special skills are required

#### How will the tool be used?

- Requires R statistical package
- Requires Windows environment
- This tool is written in C-code and can be modified to be a tool for the ALA
- Desktop application
- user input required

---

<sup>189</sup> <http://www.biomaps.net.au/gdm/>

<sup>190</sup> Ferrier and Manion (2007) GDM tutorial

<sup>191</sup> Ferrier, S., Manion, G., Elith, J. and Richardson, K. (2007) Using generalized dissimilarity modelling to analyse and predict patterns of beta diversity in regional biodiversity assessment. *Diversity and Distributions* **13**: 252-264. available at: <http://www.blackwell-synergy.com/toc/ddi/13/3>

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- Download: [http://www.biomaps.net.au/gdm/GDM\\_R\\_Distribution\\_Pack\\_V1.1.zip](http://www.biomaps.net.au/gdm/GDM_R_Distribution_Pack_V1.1.zip)
- R statistical software: <http://www.r-project.org/>
- Contact: Operation - Glenn Manion: [glenn.manion@environment.nsw.gov.au](mailto:glenn.manion@environment.nsw.gov.au)
- Contact: GDM approach – Simon Ferrier [Simon.Ferrier@csiro.au](mailto:Simon.Ferrier@csiro.au)
- Licence: Negotiate with NSW Department of Environment and Conservation

**Comments**

The GDM tool can be modified to sit within, and use the architecture of the ALA.<sup>192</sup>

---

<sup>192</sup> Simon Ferrier, January 2008

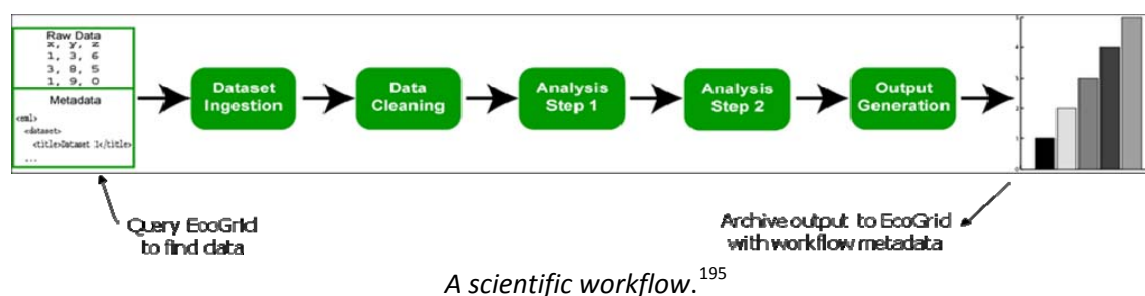
## Kepler

Summary	
Type of tool	Application
Function	Scientific workflow design
Online / Desktop	Modern desktop
Computer infrastructure	Java, platform independent
Development status	Beta version 1.0.0 beta3 January 2007
Time of use	Before ALA, during streaming, and as a post-process
Licence	Open source

**Kepler** is an open source software tool that allows scientists to design scientific workflows and execute them efficiently using emerging Grid-based approaches to distributed computation.<sup>193</sup>

### Description

Kepler is a software application for the analysis and modeling of scientific data. Using Kepler's graphical interface and components, scientists with little background in computer science can create executable models called "scientific workflow," which are a flexible tool for accessing scientific data (streaming sensor data, medical and satellite images, simulation output, observational data, etc.) and executing complex analysis on the retrieved data. Kepler is developed by a cross-project collaboration to develop open source tools to enable scientists to create and run computational experiments.<sup>194</sup>



Each workflow consists of analytical steps that may involve database access and querying, data analysis and mining, and intensive computations performed on high performance cluster computers. Each workflow step is represented by an "actor," a processing component that can be dragged and dropped into a workflow via Kepler's visual interface. Connected actors (and a few other components that we'll discuss in later sections) form a workflow, allowing scientists to inspect and display data on the fly as it is computed, make parameter changes as necessary, and re-run and reproduce experimental results.

Workflows may represent theoretical models or observational analyses; they can be simple and linear, or complex and non-linear. One of the benefits of scientific workflows is that they can be nested, meaning that a workflow can contain "sub-workflows" that perform embedded tasks.<sup>196</sup>

<sup>193</sup> <http://seek.ecoinformatics.org/Wiki.jsp?page=Kepler>

<sup>194</sup> <http://www.kepler-project.org/Wiki.jsp?page=UserFAQ>

<sup>195</sup> Periera, RS, Overview of the Science Environment for Ecological Knowledge (SEEK), 30 June 2005. <http://www.nesc.ac.uk/action/esi/download.cfm?index=2553>

<sup>196</sup> Getting Started with Kepler

from: <http://cvs.ecoinformatics.org/cvs/cvsweb.cgi/kepler-docs/user/draft-getting-started-guide.pdf>

Kepler builds upon the mature Ptolemy II framework, developed at the University of California, Berkeley.

"*Ptolemy II* is a software framework developed as part of the Ptolemy project, which studies modeling, simulation, and design of concurrent, real-time, embedded systems.<sup>197</sup>

Kepler is designed to support numerous scientific domains, including bioinformatics, ecoinformatics, geoinformatics, and others.<sup>198</sup>

Kepler includes distributed computing technologies that allow scientists to share their data and workflows with other scientists and to use data and analytical workflows from others around the world. Kepler also provides access to a continually expanding, geographically distributed set of data repositories, computing resources, and workflow libraries (e.g., ecological data from field stations, specimen data from museum collections, data from the geosciences, etc.).<sup>199</sup>

### Function

- Data analysis tools
  - Data cleaning
  - Data mining
- Analysis tools
  - Simple
  - Complex

### Why use this tool?

Kepler users with little background in computer science can create workflows with standard components, or modify existing workflows to suit their needs. Quantitative analysts can use the visual interface to create and share R and other statistical analyses. Users need not know how to program in R in order to take advantage of its powerful analytical features; pre-programmed Kepler components can simply be dragged into a visually represented workflow.<sup>200</sup>

### Who will use this tool?

- Data creators
- Data capture
- Data users

### How will the tool be used?

- Modern desktop
- Java 1.4 or later required
- Platform independent
- Can use high performance clusters
- User input required
- Supports EcoGrid access

### Where in the data chain could this tool be used?

- Data source

---

<sup>197</sup> <http://www.kepler-project.org/Wiki.jsp?page=UserFAQ>

<sup>198</sup> <http://www.kepler-project.org/Wiki.jsp?page=ProjectFAQ>

<sup>199</sup> Getting Started with Kepler,  
from: <http://cvs.ecoinformatics.org/cvs/cvsweb.cgi/kepler-docs/user/draft-getting-started-guide.pdf>

<sup>200</sup> Getting Started with Kepler  
from: <http://cvs.ecoinformatics.org/cvs/cvsweb.cgi/kepler-docs/user/draft-getting-started-guide.pdf>

- User's machine

**When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request
- As a post process, after data is with the user

**Availability**

- Kepler: <http://kepler-project.org/>
- Download: <http://www.kepler-project.org/Wiki.jsp?page=Downloads>
- Download size: Approximately 100 MB
- Installer available for Mac, Windows and Linux
- Beta version only: 1.0.0 beta 3 January 2007
- User's discussion list:  
<http://mercury.nceas.ucsb.edu/ecoinformatics/mailman/listinfo/kepler-users>
- FAQ: <http://www.kepler-project.org/Wiki.jsp?page=ProjectFAQ>
- Documentation:  
<http://kepler-project.org/Wiki.jsp?page=Publications>  
<http://www.kepler-project.org/Wiki.jsp?page=Documentation>
- Licence: Open Source see: <http://www.opensource.org/>
- Free to download

**Comments**

What are the data requirements?

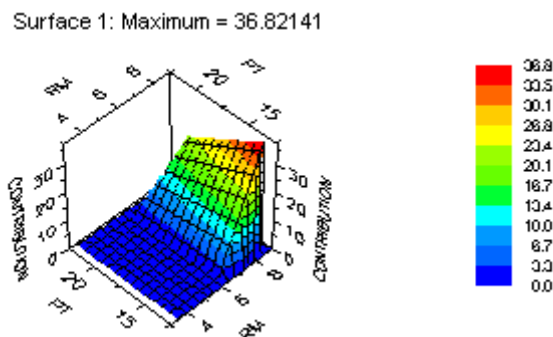
What ALA architecture is required to support the tool?

## MARS – Multivariate Adaptive Regression Splines

Summary	
Type of tool	Application
Function	Data mining, species modelling
Online / Desktop	Desktop
Computer infrastructure	Windows, Linux, R
Development status	Commercial and recent freeware
Time of use	As a post process, after data is with the use
Licence	

**MARS** - Multivariate Adaptive Regression Splines - is a regression modelling tool, able to separate relevant from irrelevant predictor variables.

### Description



*A surface plot from Salford Systems MARS.*<sup>201</sup>

MARS is a hybrid between conventional regression and recursive partitioning methods. MARS uses piece-wise linear basis functions to define the modelled relationship. Basis functions are defined in pairs, using a knot to define inflection points, and coefficients to quantify the slopes of the non-zero sections. More than one knot (i.e. more than one pair of basis functions) can be specified for a predictor variable, allowing complex non-linear relationships to be fitted.

When fitting a MARS model, knots are chosen in a forward stepwise procedure. Candidate knots can be placed at any position within the range of each predictor variable to define a pair of basis functions. At each step, the model selects the knot and its corresponding pair of basis functions that give the greatest decrease in the residual sum of squares. Knot selection proceeds until some maximum model size is reached, after which a backwards-pruning procedure is applied and those basis functions that contribute least to model fit are progressively removed. At this stage, a predictor variable can be dropped from the model completely if none of its basis functions contribute meaningfully to predictive performance.

The sequence of models generated from this process is then evaluated using generalized cross-validation, and the model with the best predictive fit is selected.<sup>202</sup>

MARS was developed in the 1990s by Jerry Friedman. MARS can be implemented in the statistics software R, and is available as a commercial stand-alone application.

<sup>201</sup> <http://www.salfordsystems.com/1211.php>

<sup>202</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data *Ecography* 29: 129-151. Online appendix to this paper: E4596 <http://www.oikos.ekol.lu.se/app.html>



## Function

- Analysis tools
- User interface
  - Personal use
  - Raw data and visual presentation

## Why use this tool?

MARS excels at finding optimal variable transformations and interactions, the complex data structure that often hides in high dimensional data. In doing so, this approach to data mining uncovers data patterns and relationships that are difficult, if not impossible, for other approaches to uncover.<sup>203</sup>

## Who will use this tool?

- Data users
  - Expert
- Special skills are required

## How will the tool be used?

- Data can be imported as spreadsheet, database or SPSS files
- Windows or Linux
- Desktop application
- MARS algorithm can be used within R, or as a commercial stand-alone application
- User input required

## Where in the data chain could this tool be used?

- User's machine

## When could this tool be used?

- As a post process, after data is with the user

## Availability

### Commercial

- Salford Systems: <http://www.salfordsystems.com/mars.php>
- Evaluation version: <http://www.salfordsystems.com/marsreg.php>
- Q&A: <http://www.salfordsystems.com/122.php>
- Licence: Contact Salford Systems

### R Project for Statistical Computing

- R: <http://www.r-project.org/>
- MDA package (mixture and flexible discriminate analysis) Version 0.3-2 (Oct 2006): <http://cran.r-project.org/src/contrib/Descriptions/mda.html>
- Manual: <http://cran.r-project.org/doc/packages/mda.pdf>
- Licence: GPL-2, GNU Public Licence
- Cost: free

## Comments

---

<sup>203</sup> <http://www.salfordsystems.com/1211.php>

For species' distribution modelling, MARS has been compared to other methods.<sup>204</sup>

---

<sup>204</sup> Elith *et al* (2006) Novel methods improve prediction of species' distributions from occurrence data  
*Ecography* **29**: 129-151. available at: <http://www.blackwell-synergy.com/toc/eco/29/2>

## Marxan

Summary	
Type of tool	Application
Function	Reserve planning tool
Online / Desktop	Desktop
Computer infrastructure	Windows/DOS
Development status	Version 2.02 April 2007
Time of use	Post process
Licence	Freely available with acknowledgement

**Marxan** is a reserve planning tool.

### Description

**Marxan** is software that delivers decision support for reserve system design. Marxan finds reasonably efficient solutions to the problem of selecting a system of spatially cohesive sites that meet a suite of biodiversity targets. Given reasonably uniform data on species, habitats and/or other relevant biodiversity features and surrogates for a number of planning units (recently increased to more than 20,000) Marxan minimizes the cost (a weighted sum of area and boundary length) while meeting user-defined biodiversity targets.

The optimisation algorithm that attempts to find good systems of sites is 'simulated annealing'. The number of possible solutions is vast (for 200 planning units there are over  $1.6 \times 10^{60}$  solutions) and because the problem is NP-complete there is no possible method for extracting an optimal solution in reasonable time for large problems. Because of this there is no real hope (or indeed incentive) to find *an* optimal solution: Marxan will find good solutions using simulated annealing. The user can also invoke a variety of less sophisticated, but often faster, heuristic algorithms. We have found that one of the most useful outputs from the decision support software is the 'summed irreplaceability' output (see figure below) This output shows how often each planning unit is in one of the good systems. Planning units that are chosen more than 50% of the time can be thought of as being essential for efficiently meeting biodiversity goals. Sites that are rarely selected can be ignored.

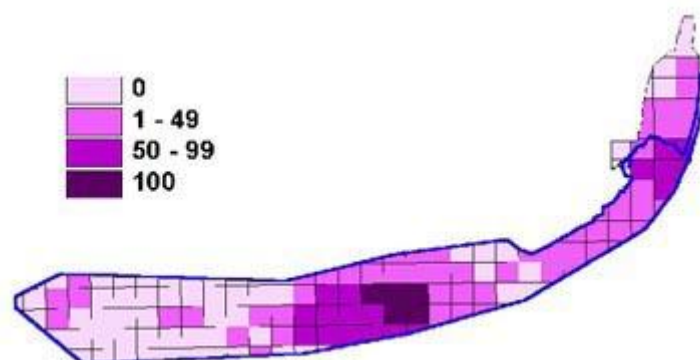


Figure. An example of SPEXAN's 'summed irreplaceability' capacity. The shading represents the number of times a 10 km x 10 km block of the Florida Keys was chosen from 100 relatively efficient solutions where the goal was representing 20% of each habitat type.<sup>205</sup>

Developed by Ian Ball and Hugh Possingham at University of Queensland.

<sup>205</sup> <http://www.ecology.uq.edu.au/marxan.htm>

### **Function**

- Analysis tools
  - Complex

### **Why use this tool?**

To provide planning information about sites that should be reserved, and those sites that may not be needed.

The basic idea is that the reserve designer has a large number of potential sites or planning units from which to select a reserve system. They wish to devise a reserve system which is made up of a selection of these planning units which will satisfy a number of ecological, social and economic criteria.<sup>206</sup>

### **Who will use this tool?**

- Data users
  - Expert
  - Skilled

### **How will the tool be used?**

- Desktop application
- Requires Windows/DOS
- Standard output is text format suitable for ArcGIS
- Standard input is a set of five text format data files
- User input is required to initially set up the data, and then running different options.

### **Where in the data chain could this tool be used?**

- User's machine
- Post data preparation

### **When could this tool be used?**

- As a post process, after data is with the user

### **Availability**

- Marxan: <http://www.ecology.uq.edu.au/index.html?page=27710>
- Download: <http://www.ecology.uq.edu.au/marxan.htm>
- Thorough manual available
- Current version 2.0.2 April 2007
- Freely available with attribution.<sup>207</sup> The right to use Marxan comes with it a responsibility to inform, share and discuss.

### **Contact:**

Professor Hugh Possingham  
Director of The Ecology Centre  
Departments of Zoology and Mathematics  
The University of Queensland  
St Lucia, QLD 4072

---

<sup>206</sup> Ball, I. R. and H. P. Possingham, (2000) MARXAN (V1.8.2): Marine Reserve Design Using Spatially Explicit Annealing, a Manual

<sup>207</sup> <http://www.ecology.uq.edu.au/index.html?page=29767&pid=29768>

AUSTRALIA  
Tel: 61 7 3365 9766  
Fax: 61 7 3365 4828  
Email: [h.possingham@uq.edu.au](mailto:h.possingham@uq.edu.au)

### Comments

- What ALA architecture is required to support the tool?

The increasing complexity of spatial planning problems has lead to the development of MarZone, a new software program that extends on the capabilities of Marxan. MarZone incorporates new functionalities to support spatial planning with multiple zones, multiple costs and multiple objectives. Using MarZone, practitioners in natural resource management can identify configurations of sites that contribute to management objectives, whilst minimising cost.<sup>208</sup>

---

<sup>208</sup> <http://www.ecology.uq.edu.au/index.html?page=69118>

## Maxent

Summary	
Type of tool	Application
Function	Modelling of species habitat
Online / Desktop	Desktop
Computer infrastructure	Platform independent. Uses Java
Development status	Version 3.1.0 November 2007
Time of use	Post process
Licence	Educational and research only

**Maxent** is a program to model species geographic distribution based on maximising entropy.

### Description

Maxent is a tool that uses a maximum-entropy approach for species habitat modeling. This software takes as input a set of layers or environmental variables (such as elevation, precipitation, etc.), as well as a set of georeferenced occurrence locations, and produces a model of the range of the given species.<sup>209</sup>

The model for a species is determined from a set of environmental or climate layers (or "coverages") for a set of grid cells in a landscape, together with a set of sample locations where the species has been observed. The model expresses the suitability of each grid cell as a function of the environmental variables at that grid cell. A high value of the function at a particular grid cell indicates that the grid cell is predicted to have suitable conditions for that species. The computed model is a probability distribution over all the grid cells. The distribution chosen is the one that has maximum entropy subject to some constraints: it must have the same expectation for each feature (derived from the environmental layers) as the average over sample locations.<sup>210</sup>

Written by Steven Phillips, Miro Dudik and Rob Schapire, with support from AT&T Labs-Research, Princeton University, and the Center for Biodiversity and Conservation, American Museum of Natural History.

### Function

- Analysis tools
  - Species geographic distribution: single algorithm applied once
- Visualisation tools
  - Maps

### Why use this tool?

To model the geographic distribution of species

### Who will use this tool?

- Data users
  - Expert
  - Interest group

---

<sup>209</sup> <http://www.cs.princeton.edu/~schapire/maxent/>

<sup>210</sup> Maxent help file

### **How will the tool be used?**

- Platform independent
- Requires Java 1.4.2 or later
- Batch mode available

Input files, output directory and algorithm parameters can be specified through the user interface, or on a command line. The user interface is best for doing single runs, while the command line is useful for repeated runs or automatically performing a sequence of runs with variations in the set of inputs.<sup>211</sup>

### **Where in the data chain could this tool be used?**

- User's machine

### **When could this tool be used?**

- As a post process, after data is with the user

### **Availability**

- Source: <http://www.cs.princeton.edu/~schapire/maxent/>
- Licence: Educational and research only  
see: <http://www.cs.princeton.edu/~schapire/maxent/>
- Cost: Free
- Version 3.1, November 2007
- No warranty

### **Comments**

- Documentation includes a tutorial, test data, help file.
- A discussion group is active.

---

<sup>211</sup> Maxent help file

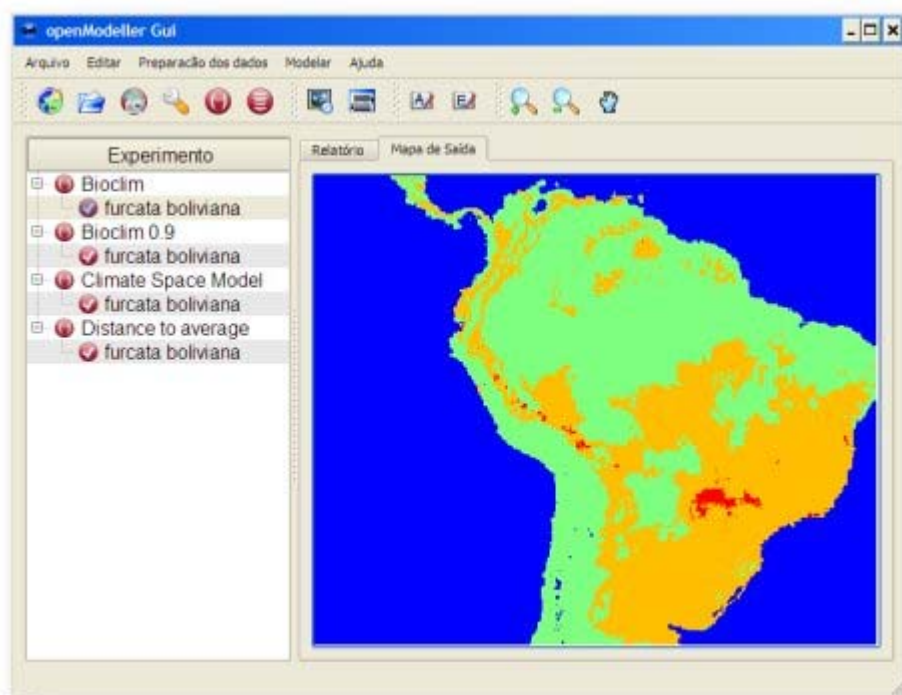
## openModeller

Summary	
Type of tool	A suite of tools
Function	Niche modelling
Online / Desktop	Desktop
Computer infrastructure	Cross-platform
Development status	Operational and active. Version 1.06 October 2007
Time of use	As a post process, when data is with user
Licence	Open Source

**openModeller** is a generic framework for carrying out fundamental niche modelling experiments - typically used to predict species distribution given a set of environmental raster layers.<sup>212</sup>

### Description

The openModeller project aims to provide a flexible, user friendly, cross-platform environment where the entire process of conducting a fundamental niche modelling experiment can be carried out. The software includes facilities for reading species occurrence and environmental data, selection of environmental layers on which the model should be based, creating a fundamental niche model and projecting the model into an environmental scenario. A number of fundamental niche modelling algorithms are provided as plug-ins, including GARP, Climate Space Model, Bioclimatic Envelopes, and others. Additional algorithms are planned for the future.<sup>213</sup>



*Desktop Map Viewer*<sup>214</sup>

*openModeller*

<sup>212</sup> <http://openmodeller.sourceforge.net/>

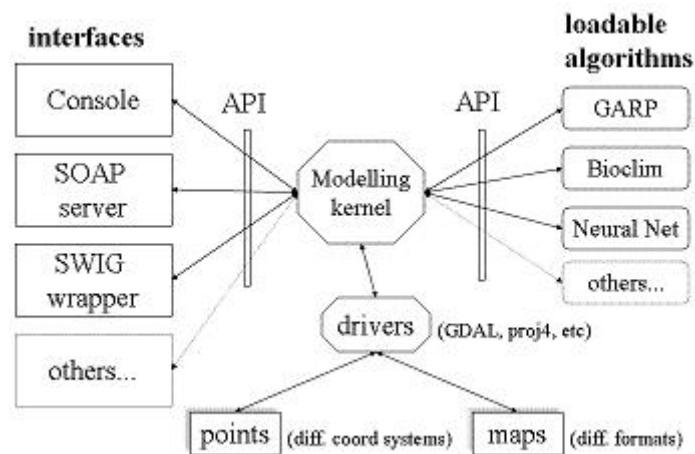
<sup>213</sup> <http://openmodeller.sourceforge.net/>

<sup>214</sup> [http://sourceforge.net/project/screenshots.php?group\\_id=101808&ssid=49152](http://sourceforge.net/project/screenshots.php?group_id=101808&ssid=49152)



openModellerDesktop is a cross platform, user friendly interface to the openModeller library. With openModellerDesktop you create an experiment consisting of one or more models. When you run the experiment, a model is computed to predict the distribution of each taxon you have selected. The model is then used to generate ('project') a probability distribution map. An embedded mapping component in openModellerDesktop allows you to visualise the probability distribution map, and shows the original occurrence points overlaid on to of the map.<sup>215</sup>

Models are generated by an algorithm that receives as input a set of occurrence points (latitude/longitude) and a set of environmental layer files. It's being written in C++ ANSI with platform independence in mind. It accepts different algorithms (now **BIOCLIM** and simple Cartesian distance algorithms, **GARP**, Climate Space Model, and in the near future **GAM**, **GLM**, **Neural Nets**, etc). It uses GDAL to read several map file formats and proj4 to convert between different georeferencing systems and projections.



A client-server architecture will be used as a first approach, enabling the existence of different client interfaces (desktop, command line and web-based). In the future some tasks could be performed in a distributed way<sup>216</sup>

The project is currently being developed by the Centro de Referência em Informação Ambiental (CRIA), Escola Politécnica da USP (Poli), and Instituto Nacional de Pesquisas Espaciais (INPE) as an open-source initiative. It is funded by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP), the Incofish project, and by individuals that have generously contributed their time. Previous collaborators include the BDWorld project, the University of Kansas Natural History Museum & Biodiversity Research Center (KU), and other individual participants.<sup>217</sup>

### Function

- Analysis tools
  - Complex – niche modelling
- User interface
  - Personal use with a visual presentation

### Why use this tool?

- For niche modelling

### Who will use this tool?

<sup>215</sup> [http://sourceforge.net/project/platformdownload.php?group\\_id=101808&sel\\_platform=4998](http://sourceforge.net/project/platformdownload.php?group_id=101808&sel_platform=4998)

<sup>216</sup> [http://openmodeller.sourceforge.net/index.php?option=com\\_content&task=view&id=2&Itemid=3](http://openmodeller.sourceforge.net/index.php?option=com_content&task=view&id=2&Itemid=3)

<sup>217</sup> <http://openmodeller.sourceforge.net/>

- Data users
  - Expert

#### **How will the tool be used?**

- This is a desktop tool which can be run with a console
- openModeller library package can be run from the command line of a scripted environment<sup>218</sup>
- Versions are available for Windows, Mac, Unix
- Source code is available

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- As a post process, after data is with the user

#### **Availability**

- openModeller: <http://openmodeller.sourceforge.net/>
- download: [http://sourceforge.net/project/platformdownload.php?group\\_id=101808&sel\\_platform=4998](http://sourceforge.net/project/platformdownload.php?group_id=101808&sel_platform=4998)
- Licence: Open Source - <http://sourceforge.net/>
- Cost: Free

#### **Comments**

---

<sup>218</sup> [http://sourceforge.net/project/platformdownload.php?group\\_id=101808&sel\\_platform=4998](http://sourceforge.net/project/platformdownload.php?group_id=101808&sel_platform=4998)

## PRIMER – Multivariate Statistics for Ecologists

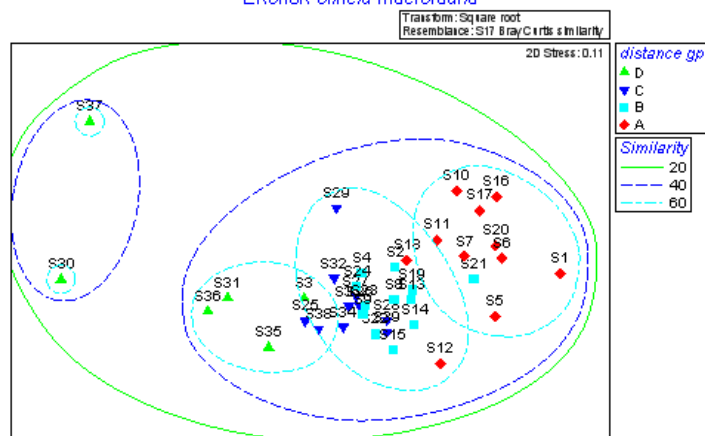
Summary	
Type of tool	Application
Function	Statistical analysis of species abundance
Online / Desktop	Desktop
Computer infrastructure	Windows, .NET
Development status	Commercial. Version 6
Time of use	After data is with user
Licence	Single user

PRIMER 6 is a collection of specialist routines for analysing species or sample abundance.<sup>219</sup>

### Description

PRIMER 6 (Plymouth Routines In Multivariate Ecological Research) consists of a wide range of univariate, graphical and multivariate routines for analysing the species/samples abundance (or biomass) matrices that arise in biological monitoring of environmental impact and more fundamental studies in community ecology, together with associated physico-chemical data. The methods make few, if any, assumptions about the form of the data<sup>220</sup>

*Ekofisk oilfield macrofauna*



PRIMER 6 multi-dimensional scaling.<sup>221</sup>

The basic routines of the package cover:

- hierarchical clustering into sample (or species) groups (CLUSTER);
- ordination by non-metric multidimensional scaling (MDS) and principal components (PCA) to summarise patterns in species composition and environmental variables;
- permutation-based hypothesis testing (ANOSIM), an analogue of univariate ANOVA which tests for differences between groups of (multivariate) samples from different times, locations, experimental treatments etc;
- identifying the species primarily providing the discrimination between two observed sample clusters (SIMPER);
- the linking of multivariate biotic patterns to suites of environmental variables (BEST);
- comparative (Mantel-type) tests on similarity matrices (RELATE);
- standard diversity indices;

<sup>219</sup> <http://en.wikipedia.org/wiki/Primer-E> Primer

<sup>220</sup> <http://www.zen87707.zen.co.uk/primer-e/primer6.htm>

<sup>221</sup> <http://www.zen87707.zen.co.uk/primer-e/primer6.htm>

- dominance plots;
- species abundance distributions;
- aggregation of arrays to allow data analysis at higher taxonomic levels, etc.

A further unique feature of PRIMER 6 is the ability to calculate biodiversity indices based on the taxonomic distinctness or relatedness of the species making up a quantitative sample or species list, indices whose statistical properties are robust to variations in sampling effort. These routines allow formal hypothesis tests for change in biodiversity structure at a location (as measured by average and variation in taxonomic 'breadth' of the species list), from that 'expected' from a larger, regional species pool. It provides a possible way of comparing biodiversity patterns over wide space and time scales, when sampling effort is not controlled.<sup>222</sup>

### **Function**

- Analysis tools
- User interface
  - Personal use
  - Visual presentation

### **Why use this tool?**

- Analysis of species abundance

### **Who will use this tool?**

- Data users
  - Expert
- Special skills are required

### **How will the tool be used?**

- Desktop application
- Windows .NET environment
- User input required

### **Where in the data chain could this tool be used?**

- User's machine

### **When could this tool be used?**

- As a post process, after data is with the user

### **Availability**

- PRIMER-E: <http://www.primer-e.com/>
- Demo: <http://www.zen87707.zen.co.uk/primer-e/demo/demo.htm>
- Contact: [admin@primer-e.com](mailto:admin@primer-e.com)
- Cost: Government licence GB£420 single use
- Version 6.1

### **Comments**

---

<sup>222</sup> <http://www.zen87707.zen.co.uk/primer-e/primer6.htm>

## SGA tool – Survey Gap Analysis

Summary	
Type of tool	C program
Function	Survey Gap Analysis modelling
Online / Desktop	Desktop
Computer infrastructure	Requires a front end to interface with other tools and users
Development status	Active
Time of use	Post process, after data is with user
Licence	Upon negotiation

**SGA** - Survey Gap Analysis tool helps to design a biodiversity survey that will best complement the existing survey effort by identifying those areas least well surveyed in terms of environmental conditions.<sup>223</sup>

### Description

The Survey Gap Analysis application enables a user to utilise locations of existing specimen records and mapped environmental variables to create a mapped complementarity surface indicating the relative priority for additional survey or collection effort throughout the region of interest (ROI). The priority being based on the potential for an area (based on climatic conditions) to complement existing survey effort in the region of interest.<sup>224</sup>

This excerpt from Funk *et al* 2005 briefly describes the Survey Gap Analysis tool:<sup>225</sup>

"... it analyses the survey coverage of a region in relation to the underlying continuous environmental and geographical space, rather than in terms of arbitrary classes. Faith & Walker's (1996)<sup>226</sup> environmental diversity (ED) measure, based on the p-median criterion, was developed for selecting sets of sites that represent regional biodiversity by providing best possible coverage of regional environmental variation. It functions by measuring how well a set of sites covers the continuous environmental space and evaluating the potential improvement that any given site would make if added. The technique, based on the finding that sampling different parts of the overall environmental space yields a good representation of the biological diversity of a region (Faith & Walker, 1996), can equally be applied to the problem of selecting survey sites."

The basic calculations can be described as follows. Given a partial set of survey sites, SGA identifies a new site that would be expected to contribute the greatest number of additional species. Note that this is not the same as finding a site that would have the greatest total number (richness) of species. Instead, it is the site with the greatest *complementarity* to the existing survey sites. Complementarity is widely used in biological conservation, and conventionally refers to some count of the number of additional species provided by a new site. In the SGA context, we cannot explicitly count such gains. Instead, these complementarity values

<sup>223</sup> <http://www.bdtracker.net/node/79>

<sup>224</sup> <http://gbifmapa.austmus.gov.au/mapa/help.jsp#WhatSGA>

<sup>225</sup> Funk, V.A, Richardson, K.S., and Ferrier, S 2005. Survey-gap analysis in expeditionary research: where do we go from here? *Biological Journal of the Linnaen Society*, 2005, 85, 549-567

available at [http://www.mnh.si.edu/biodiversity/bdg/Funk\\_Richardson\\_Ferrier.pdf](http://www.mnh.si.edu/biodiversity/bdg/Funk_Richardson_Ferrier.pdf)

and as quoted in <http://gbifmapa.austmus.gov.au/mapa/help.jsp#WhatSGA>

<sup>226</sup> Faith, D. P. & Walker, P. A. (1996) Environmental diversity: on the best-possible use of surrogate data for assessing the relative biodiversity of sets of areas. *Biodiversity and Conservation* 5, 399-415. available at: <http://www.springerlink.com/content/m11227t37q763163/?p=a60a28fad9b546e387d4792b6d5b4779&pi=0>

are estimated using ED (providing *ED-complementarity* values; Faith *et al* 2004;<sup>227</sup> Funk *et al* 2005).<sup>228</sup>

Developed by Simon Ferrier at NSW Department of Environment and conservation.

### Function

- Analysis tools
  - Complex – gap analysis
- User interface
  - The SGA tool works with raw data needing a front end to interface with users

### Why use this tool?

Biodiversity surveys are expensive undertakings requiring careful planning, and specialized resources in terms of personnel skills and equipment. It is essential then that the data obtained through new surveys complements existing data and maximizes the usefulness of the new data for conservation planning purposes. Many, if not all, museum collections are characterized by biased sampling resulting from either ad-hoc collection techniques or from planning that is based on ease of access and which only considers geographic, rather than environmental, coverage when locating survey sites. Survey Gap Analysis can be instrumental in reducing bias and thereby more effectively answering the question "If one is interested in obtaining an overall knowledge of the biodiversity (or of a taxon) of an area, and if there are insufficient data, then where should survey data be gathered?" Funk *et al* 2005.<sup>229</sup>

### Who will use this tool?

- Data users
  - Expert
- Special skills are required

### How will the tool be used?

- The SGA tool is a C code program
- Requires a front end to interface to other programs or users

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- As a post process, after data is with the user

### Availability

- Contact: Simon Ferrier [Simon.Ferrier@csiro.au](mailto:Simon.Ferrier@csiro.au)
- Licence: Negotiate with NSW Department of Environment and Conservation

### Comments

This SGA tool is used within GBIF MAPA.

---

<sup>227</sup> Faith, D. P., Ferrier, S. & Walker, P. A. (2004) The ED strategy: how species-level surrogates indicate general biodiversity patterns through an "environmental diversity" perspective. *Journal of Biogeography* **31**, 1207-1217. [http://www.amonline.net.au/systematics/pdf/jbi\\_faith\\_2004.pdf](http://www.amonline.net.au/systematics/pdf/jbi_faith_2004.pdf)

<sup>228</sup> <http://gbifmapa.austmus.gov.au/mapa/help.jsp#WhatSGA>

<sup>229</sup> <http://gbifmapa.austmus.gov.au/mapa/help.jsp#WhatSGA>

In GBIF MAPA the SGA tool produces a map showing the distribution of current ED complementarity values for all possible new sites. Darker colours indicate larger values; the red flag is the suggested choice for a new site, as it is the one sitting at the "highest peak". Using the complementarity surface as a guide the user can move the suggested site to take into account access constraints (using roads and rivers GIS layers). Once accepted as a new site the red flag turns green and cannot be moved again.

Once a site is accepted and the analysis run again to select another site, a new complementarity surface is created showing a different pattern of ED complementarity values based on the use of the original survey sites you started the analysis with and the new sites chosen by SGA. This highlights the dynamic nature of the survey gap analysis - the complementarity value of a site always depends on the set of sites already selected.<sup>230</sup>

---

<sup>230</sup> <http://gbifmapa.austmus.gov.au/mapa/help.jsp#WhatSGA>

## SPRAT – Species Profile and Threats Database

Summary	
Type of tool	Database and analysis tool
Function	Threatened species analysis
Online / Desktop	Online (for public) and (in-house) desktop
Computer infrastructure	Browser (online) and Oracle database (desktop)
Development status	Active development
Time of use	At time of user request
Licence	Contact Australian Department of Environment

A database of threatened species, with their distributions.

### Description

**SPRAT – Species Profile and Threats Database** is designed to provide information about species and ecological communities listed under the *Environment Protection and Biodiversity Conservation Act 1999*. It provides information on what the species looks like, its population and distribution, habitat, movements, feeding, reproduction and taxonomic comments. The information has been compiled by summarising information from a range of sources and contributors. At this stage profiles are not available for all species and ecological communities, but will be regularly added to the database.<sup>231</sup>

This database references collections data of over 25 million georeferenced specimens and observations. In-house tools have been developed for data cleaning and validating.<sup>232</sup>

Modelling is done with BIOCLIM and NVIS – National Vegetation Information System.

### Function

- Visualisation tools
  - Maps
- Analysis tools
  - Simple – distribution
- User interface
  - Personal use
  - Visual presentation (online) and raw data is available from the desktop version

### Why use this tool?

To determine the conservation status and likely distribution of threatened species in Australia

### Who will use this tool?

- Data users
  - Expert
  - Interest groups
  - General public

### How will the tool be used?

- Online and desktop version

<sup>231</sup> <http://www.environment.gov.au/cgi-bin/sprat/public/sprat.pl>

<sup>232</sup> Maris Ozolins, Department of Environment, January 2008



- Online is a searchable database
- Desktop version is much more sophisticated allowing species modelling
- Uses common names, scientific names, and wildcards
- User input is required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- SPRAT online: <http://www.environment.gov.au/cgi-bin/sprat/public/sprat.pl>
- Contact: Jeff Tranter [Jeff.Tranter@environment.gov.au](mailto:Jeff.Tranter@environment.gov.au) Australian Department of Environment, Water, Heritage and the Arts
- Licence: Contact Australian Department of Environment for access to this tool

**Comments**

- Restricted to those species that are endangered under the EPBC act.
- The Australian Department of Environment, Water, Heritage and the Arts are open to the concept of the ALA making use of some of their software tools.

## Triana

Summary	
Type of tool	Framework with tools
Function	Data manipulation, modelling
Online / Desktop	Desktop
Computer infrastructure	OS independent, written in Java
Development status	In production and stable
Time of use	Post process when data is with user
Licence	Open Source, Apache Software Licence

**Triana** is a distributed problem solving environment. It combines workflow with a library of data analysis tools and distributed computing services.

### Description

Triana is an open source, problem solving environment that combines an intuitive visual interface with powerful data analysis tools. Already used by scientists for a range of tasks, such as signal, text and image processing, Triana includes a large library of pre-written analysis tools and the ability for users to easily integrate their own tools.<sup>233</sup>

Triana is a graphical environment that allows you to create powerful computer programs and to use them, with a minimum of effort and no programming. Using Triana, you simply assemble your program from a set of building-blocks that you drag into a work-space window and connect up using your mouse. With a click of the mouse the program will perform whatever operations you want. You can tell Triana to execute your program just once or continuously, as long as data is available to it.

Since Triana is written in pure Java, it will run on almost any computer.

- Use Triana on a wide variety of data: numerical data, either taken from an experiment or generated by Triana; audio data; images; even text files.
- Triana comes with a wide variety of built-in tools. There is an extensive signal-analysis toolkit, an image-manipulation toolkit, a desk-top publishing toolkit, and many more.
- Triana will display your data, either in a text-editor window or in a versatile graph-display window.
- Triana is particularly good at automating repetitive tasks, such as continuously monitoring the spectrum of data that comes from an experiment that runs for days or even years.
- If the tools supplied with Triana do not do what you need, Triana contains a wizard that helps you to create new ones, with parameter windows.

Triana is being developed at Cardiff University in the UK.<sup>234</sup>

### Function

- Analysis tools
  - Simple
  - Complex

### Why use this tool?

---

<sup>233</sup> <http://www.trianacode.org/>

<sup>234</sup> Triana User Guide <https://forge.nesc.ac.uk/docman/view.php/33/104/UserGuide.pdf>

- Use a graphical framework to solve scientific problems such as with Biodiversity World.

**Who will use this tool?**

- Data users
  - Expert

**How will the tool be used?**

- Desktop
- Platform independent Java application
- User input required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- Triana: <http://www.trianacode.org/>
- Download: <https://forge.nesc.ac.uk/projects/triana/>
- User guide: [https://forge.nesc.ac.uk/docman/?group\\_id=33](https://forge.nesc.ac.uk/docman/?group_id=33)
- Version 3.2.3 April 2007
- Licence: Apache Software Licence
- Cost: Free to download and use
- Support: mailing lists, developer site

**Comments**

Triana can use the tools of Biodiversity World (if it exists).

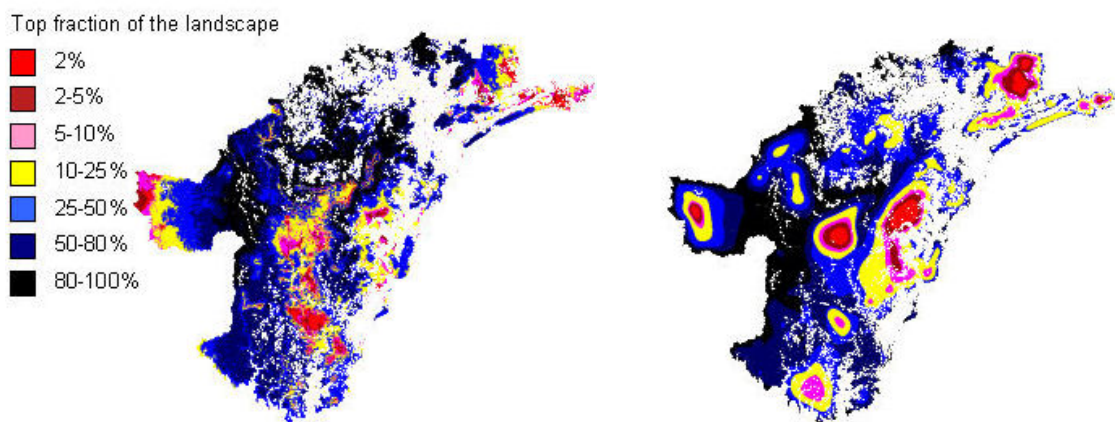
## Zonation

Summary	
Type of tool	Application
Function	Conservation planning
Online / Desktop	Online
Computer infrastructure	Windows
Development status	Advanced development. Version 1.01, December 2006
Time of use	Post process
Licence	Freely usable for non-commercial uses

**Zonation** is a reserve selection framework for spatial conservation planning. It identifies areas important for retaining habitat quality and connectivity for multiple species, indirectly aiming at species' long-term persistence. Zonation can be used for various purposes such as spatial conservation prioritization, conservation assessment, reserve selection and reserve network design.<sup>235</sup>

### Description

Zonation produces a hierarchical prioritization of the landscape based on the conservation value of sites (cells), iteratively removing the least valuable cell (accounting for complementary) from the landscape until no cells remain. In this way, landscapes can be zoned according to their value for conservation.<sup>236</sup>



*During each computation the Zonation produces detailed information about the decrease in species distributions as landscape is lost. The example area here is a 120 x 140 km<sup>2</sup> region in eastern Australia. The data includes 7 priority fauna species and a 649 x 555 grid at 200 m cell resolution<sup>237</sup>.*

The program produces, among other things, basic raster files from each run, which can be imported to GIS software to create maps or to conduct further analyses. The data requirements for the program are realistic and it can be run with large datasets containing up to 2000 species or 16 million element landscapes on an ordinary desktop PC.<sup>238</sup>

<sup>235</sup> <http://www.helsinki.fi/bioscience/consplan/>

<sup>236</sup> <http://www.helsinki.fi/bioscience/consplan/Introduction.html>

<sup>237</sup> <http://www.helsinki.fi/bioscience/consplan/Introduction.html>

<sup>238</sup> <http://www.helsinki.fi/bioscience/consplan/Introduction.html>

Zonation includes several practical features such as aggregation methods, uncertainty analysis, species prioritization and replacement cost analysis for current or proposed reserves. The data requirements are realistic and Zonation can analyse relatively large data sets in reasonable time. Windows software package includes user manual and tutorial.<sup>239</sup>

## **Components**<sup>240</sup>

### **Analyses**

- Identification of optimal reserve areas
- Planning methods
  - Core-area Zonation
  - Additive benefit function
  - Targeting benefit function

### **Data**

- Large-scale grids with
  - presence/absence -data
  - probabilities of occurrence
  - abundance/density -data
- Cost and mask layers

### **Features**

- Connectivity methods
  - Distribution Smoothing
  - Boundary Quality Penalty
  - Boundary Length Penalty
- Uncertainty analysis aiming at reliable conservation decisions
- Species prioritization
- Clearly defined tradeoffs between species

The software has been written by Atte Moilanen of the Metapopulation Research Group, University of Helsinki.<sup>241</sup>

## **Function**

- Analysis tools
  - Complex – gap analysis: recursive
- User interface
  - Personal use
  - Visual presentation

## **Why use this tool?**

Zonation links species distribution modelling directly to quantitative reserve planning. Zonation includes species-specific handling of connectivity and natural weighting of species. It can analyse relatively large data sets in reasonable time. The analysis is deterministic and the results of a Zonation run can be summarized in a single graph with zones.

---

<sup>239</sup> <http://www.helsinki.fi/science/metapop/english/Software.htm>

<sup>240</sup> <http://www.helsinki.fi/bioscience/consplan/Components.html>

<sup>241</sup> <http://www.helsinki.fi/science/metapop/english/Software.htm>

Zonation includes a number of useful novel features, such as distribution smoothing and boundary quality penalty aggregation methods as well as the uncertainty analysis. Core-area Zonation and additive benefit function Zonation are also novel planning options.<sup>242</sup>

#### Who will use this tool?

- Data users
  - Expert

#### How will the tool be used?

Typically one would enter one grid per species. Each cell would have either an observation of population size at that location, or, more commonly, a probability of occurrence or abundance predicted using a statistical habitat model.<sup>243</sup>

**What limitations does the software have?** Data sets that have been run with Zonation include 700 spp ´ 1M informative elements and 122 spp ´ 2M elements in the grid. Maximum data is limited by computer memory, and on a PC with 4GB of memory approximately 700 spp ´ 1M elements (or equivalently 70 spp ´ 10M) is close to the limit.<sup>244</sup>

**What does Zonation presently not do?** It does not work with vector-based planning units. It only allows a limited set of interactive planning analyses. It cannot handle multiple alternative land-use options. Zonation is for implement-in-one-go planning instead of multi-year incremental design of reserves.<sup>245</sup>

- Desktop application
- Windows OS
- User input is required

#### Where in the data chain could this tool be used?

- User's machine

#### When could this tool be used?

- As a post process, after data is with the user

#### Availability

- Zonation: <http://www.helsinki.fi/bioscience/consplan/index.html>
- Download: <http://www.helsinki.fi/bioscience/consplan/>
- User manual: [http://www.helsinki.fi/bioscience/consplan/Zonation\\_UserManual.pdf](http://www.helsinki.fi/bioscience/consplan/Zonation_UserManual.pdf)
- Flyer: [http://www.helsinki.fi/bioscience/consplan/Zonation\\_Leaflet.pdf](http://www.helsinki.fi/bioscience/consplan/Zonation_Leaflet.pdf)
- Version 1.01, December 2006
- Cost: Free. Freely usable for non-commercial uses
- Copyright: Atte Moilanen, University of Helsinki, Finland.
- No warranty. Use at own risk.

#### Comments

Zonation may deal better with species data than Marxan.

---

<sup>242</sup> <http://www.helsinki.fi/bioscience/consplan/Questions.html>

<sup>243</sup> <http://www.helsinki.fi/bioscience/consplan/Questions.html>

<sup>244</sup> <http://www.helsinki.fi/bioscience/consplan/Questions.html>

<sup>245</sup> <http://www.helsinki.fi/bioscience/consplan/Questions.html>

## Taxonomy

## Anatomical Atlas of Flies

Summary	
Type of tool	Illustration of fly anatomy
Function	Teaching aid, taxonomy, identification
Online / Desktop	Online
Computer infrastructure	Web browser, broadband
Development status	Complete
Time of use	At the time of user request
Licence	CSIRO copyright

The **Anatomical Atlas of Flies** can be used as a teaching aid or in conjunction with any fly key to find out the name of fly parts.<sup>246</sup>

### Description



The Anatomical Atlas of Flies can be used as a standalone resource to accompany any fly key or as an aid for teaching fly anatomy. The atlas works both ways: users can either click on a part to discover its name, or click on a name to discover the location and shape of a part. Common synonyms for anatomical terms are available from the information button that appears when terms and structures are highlighted.

The atlas relies on high resolution digital images of flies, and allows the user to change magnifications to see fine detail. The Atlas uses examples from the 4 major fly groups, representing different anatomical expressions of flyness. The user can learn fly anatomy by examining all the different parts of a single fly or highlight a particular structure and navigate between flies. In this way the user can recognize how the structure changes shape and form across this vast taxonomic spectrum. This is a graphical representation of transformational homology.

The Anatomical Atlas was created by CSIRO Entomology to accompany an ABRIS-funded identification key to fly families of Australia and US NSF-funded research into the evolutionary history of flies.

Produced by Anne Hastings, David Yeates and Joanna Hamilton (CSIRO Entomology). The 'Atlas' was made with Macromedia Flash, Adobe Photoshop and a high resolution digital camera mounted on a stereo microscope.<sup>248</sup>

### Function

<sup>246</sup> <http://www.csiro.au/resources/ps252.html>

<sup>247</sup> <http://www.csiro.au/resources/ps252.html>

<sup>248</sup> <http://www.ento.csiro.au/biology/fly/fly.html>



- Visualisation tools
  - Images
- Taxonomy
- User interface
  - Personal
  - Visual presentation

#### **Why use this tool?**

- To investigate the anatomy of the fly
- To accompany any fly key

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data users
  - Expert
  - Interest groups
  - General public

#### **How will the tool be used?**

- Online tool requires broadband
- User input is required

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- At the time of a user request
- As a post process, after data is with the user

#### **Availability**

- Anatomical Atlas of Flies: <http://www.ento.csiro.au/biology/fly/fly.html>
- CSIRO information sheet: <http://www.csiro.au/resources/ps252.html>
- Also available with the Lucid key on CD - *On The Fly: the interactive atlas and key to Australian fly families*- see <http://www.csiro.au/resources/ps236.html>
- Contact: David Yeates [David.Yeates@csiro.au](mailto:David.Yeates@csiro.au),
- Cost: free
- Copyright CSIRO 2004 <http://www.csiro.au/org/CopyrightNotice.html>

#### **Comments**

## APII – Australian Plant Image Index

Summary	
Type of tool	Database of plant images
Function	Taxonomy, images
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Stable
Time of use	Before data is with ALA, At time of user request.
Licence	Personal and non-commercial use, with attribution

The **Australian Plant Image Index** is the National collection of photographs and illustrations of Australia's flora.<sup>249</sup>

### Description



*Xanthorrhoea macronema*. Photo by G Butler<sup>250</sup>

The Australian National Botanic Gardens (ANBG) has for many years been building up a collection of photographs and illustrations of Australia's native plants.

Originally these were photos taken on field collecting trips to supplement the herbarium specimens and data associated with the living plants. In recent years many other photographers have contributed to the Index. We are now embarking on an active plan to encourage people to contribute reliably identified photographs to the Index.

<sup>249</sup> <http://www.anbg.gov.au/anbg/photo-collection/photo.index.html>

<sup>250</sup> <http://www.anbg.gov.au/photo/apii/id/a/4509>

Our aim is to have a good quality coloured slide of every Australian native plant represented in the Australian Plant Image Index.

Portraits of plant flowers form the bulk of the Index, but photographs of fruit, bark, shrubs or whole trees are also included. A collection of other illustrative material is also catalogued.

Photographs in the Index are taken as 35 mm colour slides.

The most significant aspect of the Index, and that which will separate it from many other photo libraries, is our emphasis on accurate botanical naming of the plants.<sup>251</sup>

An internet gateway to the database catalogue of the photograph collection is available, as well as gateways to related information. Only a small percentage of those photographs in the collection have been digitized, they are available at low resolution on the web.<sup>252</sup>

### **Function**

- Visualisation tools
  - Images
- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Visual presentation

### **Why use this tool?**

- Plant images

### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data users
  - Expert
  - Interest groups
  - General public
- No special skills required

### **How will the tool be used?**

- Online database
- An internet gateway to the database catalogue of the photograph collection is available, as well as gateways to related information

Search can be made by:<sup>253</sup>

- genus and species
- plant family
- simple or complex search criteria
- lists of plant genera or plant family

---

<sup>251</sup> <http://www.anbg.gov.au/anbg/photo-collection/photo.index.html>

<sup>252</sup> <http://www.anbg.gov.au/anbg/photo-collection/index.html>

<sup>253</sup> <http://www.anbg.gov.au/anbg/photo-collection/index.html>

- non-plant photos
- line illustrations for selected genus

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- At the time of a user request

**Availability**

- API: <http://www.anbg.gov.au/anbg/photo-collection/photo.index.html>
- Search: <http://www.anbg.gov.au/anbg/photo-collection/index.html>
- Contact: The Photograph Curator, ANBG, [photo@anbg.gov.au](mailto:photo@anbg.gov.au)
- Cost: free for government web pages; reproduction costs apply for other uses  
<http://www.anbg.gov.au/anbg/photo-collection/photo-collection-use.html>
- Copyright: Images are copyright. Some personal and non-commercial use is permitted with attribution and linking to ANBG.  
<http://www.anbg.gov.au/anbg/photo-collection/photo-collection-use.html>

**Comments**

## APPD – Australian Plant Pest Database

Summary	
Type of tool	Database of plant pests
Function	Pest identification
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Stable and progressive
Time of use	At time of user request
Licence	Restricted

The **Australian Plant Pest Database (APPD)** is a nationally coordinated database of plant pests and diseases.<sup>254</sup>

### Description

The Australian Plant Pest Database (APPD) provides a single point of access to existing electronic records of voucher specimens contained in databases across Australia. The system provides critical support to decision making during emergency management of incursions by exotic species. It also provides a powerful tool to assist bids for market access and to justify measures to exclude potentially harmful, exotic organisms.<sup>255</sup>

APPD - query form

This form allows users to search for records from one or more collection databases at once. A number of [help](#) pages are available to assist users with their search queries.

**User guide**  
A short user guide is now available in PDF format to assist first time APPD users.  
[Download pdf file](#) - approx 180 Kb

**Frequent users - remember to bookmark this page (press CTRL-D)**

**LATEST NEWS ON APPD**

**Select data source:**  
Click checkboxes to select data sources. Links provide additional information on individual data sources.

Arthropods:  ANIC  ASCU  BSES\*  FCNI\*  ICDB  NTEIC\*  QDDIF   
 TFIC  TPPD  UQIC  VAIC  WACALM  WINC^

Microorganisms:  APDD  WAC   
 \* deployed in Canberra. ^ currently unavailable.

**Arthropod/nematode/pathogen information:**  
 Scientific name:\*   
 Order: equals   
 Family: equals    
 Genus: equals    
 Species: equals    
 Intraspecific taxa: equals   
 Common name:\* equals

**Host information:**  
 Scientific name:\*   
 Family: equals   
 Genus: equals   
 Species: equals   
 Intraspecific taxa: equals   
 Common name:\* equals   
 Substrate:\* equals

**Coordinates:** [map input](#)  
 Latitude:  Longitude:   
 Upper right:    
 Lower left:

**Location:**  
 State: All states  Town: equals   
**Collection date:**  
 From:  (yyyy) - To:  (yyyy)

APPD query form.<sup>256</sup>

<sup>254</sup> [http://www.planthealthaustralia.com.au/our\\_projects/display\\_project.asp?category=4&ID=1](http://www.planthealthaustralia.com.au/our_projects/display_project.asp?category=4&ID=1)

<sup>255</sup> <http://www.planthealthaustralia.com.au/APPD/help/index.asp>

<sup>256</sup> Ian Naumann, Emma Lumb, Kerry Taylor, Robert Power, David Ratcliffe, and Michael Kearney (2006). The Australian Plant Pest Database: a national resource for an expert community. In Andrew Treloar, editor, *Proceedings of The Twelfth Australasian World Wide Web Conference, (AusWeb06)*, Noosa, Australia, 30th June to 4th July 2006. <http://ausweb.scu.edu.au/aw06/index.html>

The APPD query form contains three sections: <sup>257</sup>

- data sources to query
- pest and host details
- further query refinements (pest location and date specimen collected)

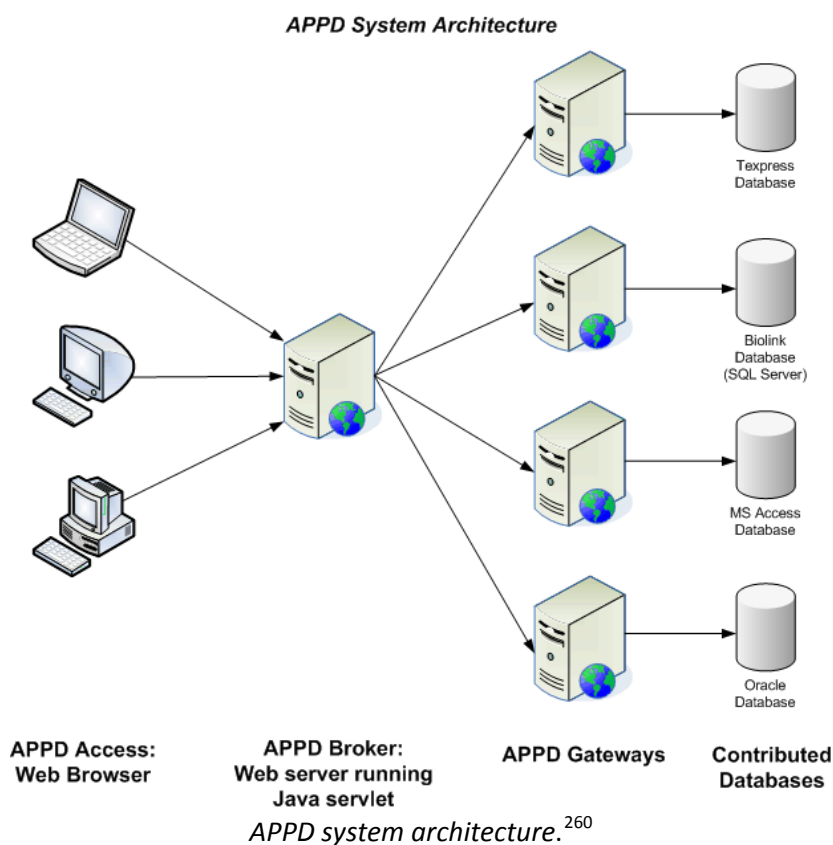
The pest scientific or common name may be used. The host scientific, common name or substrate (for non-living hosts) may be used. Scientific names may include any combinations of one or more taxonomic levels of order/family/genus/species/infraspecific taxa. <sup>258</sup>

The results can be constrained spatially by:

- Collection coordinates
- Location details of State and town
- Year of specimen collection

Results are presented on a scaleable map, and as a dataset.

Because pest information may have an impact on market access and trade, access to the APPD is restricted. Generally, only plant health or quarantine specialists and collection curators from contributing agencies have access to the database. <sup>259</sup>



<sup>257</sup> <http://www.planthealthaustralia.com.au/APPD/help/instruct-query.asp>

<sup>258</sup> <http://www.planthealthaustralia.com.au/APPD/help/instruct-query.asp>

<sup>259</sup> <http://www.planthealthaustralia.com.au/APPD/guidelines/index.asp>

<sup>260</sup> Ian Naumann, Emma Lumb, Kerry Taylor, Robert Power, David Ratcliffe, and Michael Kearney (2006). The Australian Plant Pest Database: a national resource for an expert community. In Andrew Treloar, editor, *Proceedings of The Twelfth Australasian World Wide Web Conference, (AusWeb06)*, Noosa, Australia, 30th June to 4th July 2006. <http://ausweb.scu.edu.au/aw06/index.html>

The APPD was launched in April 2002 and provides access to 14 pest collections (through 12 agencies), totalling in excess of 1 million records. This project has received funding from an Australian Government Budget Initiative administered by the Department of Agriculture Fisheries and Forestry.<sup>261</sup>

### Function

- Taxonomy
  - Identification tools
- User interface
  - Personal
  - Raw data and visual presentation

### Why use this tool?

- What is the purpose of the tool?
- Are there alternatives?

### Who will use this tool?

- Data users
  - Expert
- Semi-skilled operation

### How will the tool be used?

- Online access to this database
- User input required

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- At the time of a user request

### Availability

- APPD: <http://www.planthealthaustralia.com.au/APPD/>
- Plant Health Australia: <http://www.planthealthaustralia.com.au/site/index.asp>
- Help: <http://www.planthealthaustralia.com.au/APPD/help/index.asp>
- Restrictions: Access only available to selected specialists

### Comments

- Allows both common names and scientific names.

See paper discussing APPD architecture:

Ian Naumann, Emma Lumb, Kerry Taylor, Robert Power, David Ratcliffe, and Michael Kearney (2006). The Australian Plant Pest Database: a national resource for an expert community. In Andrew Treloar, editor, *Proceedings of The Twelfth Australasian World Wide Web Conference, (AusWeb06)*, Noosa, Australia, 30th June to 4th July 2006.

<http://ausweb.scu.edu.au/aw06/index.html>

For international commercial reasons this database has highly restricted access. It is included here as an example of a pest database; how the ALA could service this community by

---

<sup>261</sup> [http://www.planthealthaustralia.com.au/our\\_projects/display\\_project.asp?category=4&ID=1](http://www.planthealthaustralia.com.au/our_projects/display_project.asp?category=4&ID=1)

contributing tools and datasets; and how this community may be able to allow some form of restricted access to the ALA.

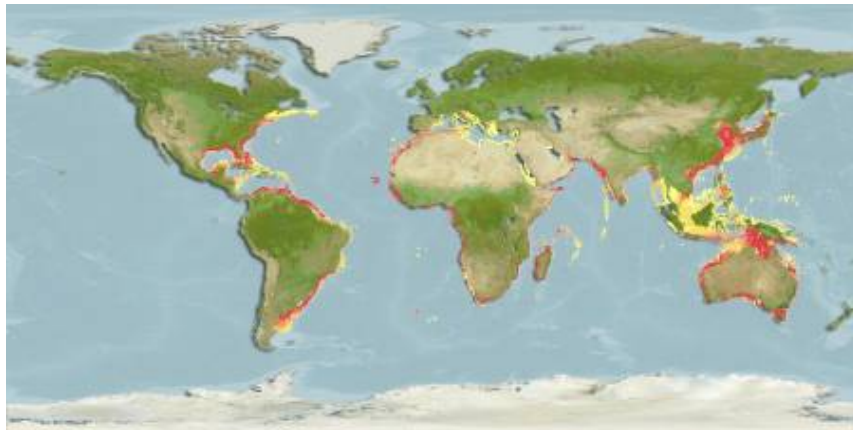


## AquaMaps

Summary	
Type of tool	Database
Function	Marine species' distribution
Online / Desktop	Online
Computer infrastructure	Browser. Data in .csv form
Development status	Developed and operational. Version 12/2007
Time of use	At time of user request
Licence	

**AquaMaps** combines numerical algorithms with expert knowledge<sup>262</sup> to create standardised distribution maps for currently 7,200 species of fishes, marine mammals and invertebrates.<sup>263</sup>

### Description



Reviewed native range of the Grey nurse shark **Carcharias Taurus**.<sup>264</sup>

AquaMaps is an approach to generating model-based, large-scale predictions of currently known natural occurrence of marine species. Models are constructed from estimates of the environmental tolerance of a given species with respect to depth, salinity, temperature, primary productivity, and its association with sea ice or coastal areas. Maps show the color-coded relative likelihood of a species to occur in a global grid of half-degree latitude / longitude cell dimensions, which corresponds to a side length of about 50 km near the equator. Predictions are generated by matching habitat usage of species, termed environmental envelopes, against local environmental conditions to determine the relative suitability of specific geographic areas for a given species. Knowledge of species' distributions within FAO areas or bounding boxes is also used to exclude potentially suitable habitat in which the species is not known to occur.

AquaMaps was originally developed by Kristin Kaschner and colleagues to predict global distributions of marine mammals (Kaschner *et al* 2006)<sup>265</sup>. The approach was based on incorporating expert knowledge into an environmental envelope or ecological niche model. The

<sup>262</sup> K. Kaschner, J. Ready, E. Agbayani, P. Eastwood, T. Rees, K. Reyes, J. Rius & R. Froese (2007) About AquaMaps: Creating standardized range maps of marine species.

See <http://fishbase.sinica.edu.tw/tools/AquaMaps/AboutAquaMaps.doc>

<sup>263</sup> <http://www.aquamaps.org/>

<sup>264</sup> [http://fishbase.sinica.edu.tw/tools/aquamaps/premap.php?expert\\_id=16&SpecID=Fis-29388&cache=1](http://fishbase.sinica.edu.tw/tools/aquamaps/premap.php?expert_id=16&SpecID=Fis-29388&cache=1)

<sup>265</sup> Kaschner, K., R. Watson, *et al* (2006). "Mapping worldwide distributions of marine mammals using a Relative Environmental Suitability (RES) model." Marine Ecology Progress Series 316: 285-310.

use of expert knowledge compensated for the effects of species misidentifications, effort biases, and the non-representative coverage of large-scale species' distributions. Such data gaps and problems are widespread in publicly available occurrence data sets that are compiled from different sources.

The approach developed for marine mammals was subsequently modified in collaboration with FishBase staff to make it more suitable for a greater range of marine organisms and to make use of data and information available in FishBase and OBIS/GBIF online databases. Display of the maps on the web has been facilitated by the use of C-squares Mapper developed by Tony Rees of CSIRO, Australia.<sup>266</sup>

### Function

- Visualisation tools
  - Maps
- Taxonomy
  - Identification tools, keys
- User interface
  - Personal
  - Raw data and visual presentation

### Why use this tool?

The AquaMaps approach was developed specifically to deal with the problems encountered when attempting to map large-scale species distributions based on existing but fragmented and potentially non-representative occurrence data. The basic and novel idea behind the AquaMaps concepts is to supplement occurrence records with independent knowledge about species distributions and habitat usage to correct for existing biases.

For instance, knowledge of the geographic extents of commercial species available from FAO can be used to define latitudinal and longitudinal bounding boxes to delimit predictions to areas known to be utilized. Area restrictions can also serve as a quality control mechanism, as they filter out outliers in occurrence records that may present misidentifications. Published depth ranges can also be used to better define associations between species distributions and depth, as these are generally based on more rigorous analyses and representative data sets, thus reducing the impacts of sampling biases.

The strength of AquaMaps compared to other species distribution modelling algorithms thus lies in its transparency and also its ability to incorporate expert knowledge and general information on species habitat usage and occurrence.<sup>267</sup>

### Who will use this tool?

- Data users
  - Expert
  - Interest groups

### How will the tool be used?

---

<sup>266</sup> K. Kaschner, J. Ready, E. Agbayani, P. Eastwood, T. Rees, K. Reyes, J. Rius & R. Froese (2007) About AquaMaps: Creating standardized range maps of marine species.

See <http://fishbase.sinica.edu.tw/tools/AquaMaps/AboutAquaMaps.doc>

<sup>267</sup> K. Kaschner, J. Ready, E. Agbayani, P. Eastwood, T. Rees, K. Reyes, J. Rius & R. Froese (2007) About AquaMaps: Creating standardized range maps of marine species.

See <http://fishbase.sinica.edu.tw/tools/AquaMaps/AboutAquaMaps.doc>

- Online application
- Data can be displayed as a jpg map showing range, suitable habitat or point data
- Data can be downloaded as a .csv file
- Feedback is available as an email link
- User input required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- At the time of a user request
- As a post process, after data is with the user

**Availability**

- AquaMaps: <http://www.aquamaps.org/>
- Documentation: <http://fishbase.sinica.edu.tw/tools/AquaMaps/AboutAquaMaps.doc>
- Cost: free
- Version 12/2007
- Licence not mentioned

**Comments**

- AquaMaps displays range data as well as point data
- The ALA could use AquaMaps as an example of mapping range data, using modelled data and expert knowledge. Or, it could create a link to AquaMaps on a species page.

## AVH – Australia’s Virtual Herbarium

Summary	
Type of tool	Distributed databases
Function	Taxonomy, species distribution
Online / Desktop	Online
Computer infrastructure	Browser, HISPID, HTML, CSV or text
Development status	Well developed
Time of use	Before data is made available to ALA. At time of user request.
Licence	Personal and non-commercial use

**Australia's Virtual Herbarium (AVH)** is an on-line botanical information resource, providing access to the wealth of data associated with scientific plant specimens in each Australian herbarium.<sup>268</sup>

### Description



By consolidating the combined data of Australia’s herbaria, the AVH can show a nation-wide view of the botanical information. Most data related to specimens are stored by the custodial institution, and there will be some resources, such as the scientific names database (Australian Plant Names Index, APNI) which is common to all.

Australia's Virtual Herbarium provides access to species pages for plants, algae and fungi of Australia.

The AVH is a collaborative project of Australian State, Territory and Commonwealth herbaria through the Council of Heads of Australian Herbaria (CHAH).

### Function

- Visualisation tools
  - Maps
  - Images
- Taxonomy
  - Identification tools
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

<sup>268</sup> <http://www.anbg.gov.au/chah/avh/avh.html>

- Easy and quick access to plant information
- Location and mapping of rare and threatened species
- Correct naming of plants for science, conservation, medicine, agriculture and horticulture
- Historical and baseline information for revegetation projects
- Inspiration and enjoyment of Australia's native vegetation
- Biological information for seed collection and fauna management
- Early notification of weed species
- Better understanding of our natural heritage
- Bioprospecting and commercial utilisation of native flora, bushfoods and natural medicines
- Selection of suitable conservation reserves
- Prioritisation of resources for collecting and biodiversity surveys
- Informed environmental planning
- Historical and heritage information <sup>269</sup>

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators
- Data users
  - Expert
  - Interest groups
  - General public
- Some botanical skills are required

#### **How will the tool be used?**

- Online search tool
- Requires accurate scientific name – genus and species
- Herbarium specimens can be displayed on a map
- Data format: HISPID, HTML, CSV or text
- Data delivery: emailed, displayed in browser, mapped or downloadable file

#### **Where in the data chain could this tool be used?**

- Data source
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request

#### **Availability**

- AVH: <http://www.anbg.gov.au/avh/>
- Help: <http://www.anbg.gov.au/chah/avh/help/index.html>
- Cost: free
- Copyright: personal non-commercial use. Images can be used on non-commercial websites  
<http://www.anbg.gov.au/copyright.html>

#### **Comments**

---

<sup>269</sup> <http://www.anbg.gov.au/chah/avh/avh.html>

- The AVH has similarities to OZCAM – Online Zoological Collections of Australian Museums.
- The ALA may use some of the concepts of AVH
- Species pages are not available for many plants

## DELTA - Intkey

Summary	
Type of tool	Set of programs
Function	Taxonomy and species identification
Online / Desktop	Desktop
Computer infrastructure	Windows
Development status	Dated. Version 5.1.1, 2000
Time of use	Before data is made available to ALA. As a post process.
Licence	Free for non-commercial use



A system of programs for computer generating descriptions and keys.

### Description

The DELTA format (DEscription Language for TAXonomy) is a flexible method for encoding taxonomic descriptions for computer processing. DELTA-format data can be used to produce natural-language descriptions, conventional or interactive keys, cladistic or phenetic classifications, and information-retrieval systems.<sup>270</sup>

The DELTA system is an integrated set of programs based on the DELTA format, includes:<sup>271</sup>

**Intkey** - a program for interactive identification and information retrieval

**DELTA Editor** - a program for creating and editing DELTA data

**Confor** - a program for translating DELTA data into other formats

The DELTA System is capable of producing high-quality printed descriptions. The most important or diagnostic attributes (derived automatically or manually) for each taxon can be emphasized in full descriptions, or short descriptions containing only these attributes can be produced. The descriptions can be fully typeset without the requirement for any manual editing.

The System was developed in the CSIRO Division of Entomology during the period 1971 to 2000. It is in use worldwide for diverse kinds of organisms, including viruses, corals, crustaceans, insects, fish, fungi, plants, and wood. It was adopted as a standard for data exchange by Biodiversity Information Standards (TDWG). The programs are free for non-commercial use.<sup>272</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Raw data and visual presentation

### Why use this tool?

- For creating species descriptions and keys
- For species identification

---

<sup>270</sup> <http://delta-intkey.com/>

<sup>271</sup> <http://delta-intkey.com/www/install.htm>

<sup>272</sup> <http://delta-intkey.com/www/overview.htm>

### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data providers
  - Institutions
- Data users
  - Expert
  - Interest groups

### **How will the tool be used?**

- Desktop application
- Windows only
- User input required

### **Where in the data chain could this tool be used?**

- Data source
- User's machine

### **When could this tool be used?**

- Before data is made available to ALA
- As a post process, after data is with the user

### **Availability**

- DELTA: <http://delta-intkey.com/>
- Overview: <http://delta-intkey.com/www/overview.htm>
- Downloads: <http://delta-intkey.com/www/programs.htm>
- Mailing list: <https://listserv.surfnet.nl/archives/delta-l.html>
- Documentation: <http://delta-intkey.com/www/programs.htm>
- Contact: MJ Dallwitz <http://delta-intkey.com/contact/dallwitz.htm>
- Version 5.1.1, 2000
- Cost: free for non-commercial use

### **Comments**

This program obviously works under modern Windows systems (XP, Vista), and the discussion list is current, though there does not appear to be any upgrades to the software since 2000.

I have found a deep concern amongst taxonomists about the lack of commitment to the ongoing viability of DELTA. Investing time and effort to create dichotomous keys using a deteriorating system is risky.



## FishBase

Summary	
Type of tool	Database
Function	Encyclopaedia of fish biology
Online / Desktop	Online and CD on desktop
Computer infrastructure	Online: browser. Desktop: Windows
Development status	Well developed since 1989
Time of use	Before data is made available to the ALA. At time of user request
Licence	Copyright lies with data contributors

**FishBase** is a database containing all the fishes known to science, with information useful for a wide range of users from fisheries managers to biology teachers.<sup>273</sup>

### Description



Bar-tailed flathead *Platycephalus endrachtensis*.<sup>274</sup>

FishBase is an information system with key data on the biology of all fishes. It includes descriptions of over 30,000 species, over 260,000 common names in hundreds of languages, over 45,000 pictures, and references to more than 40,000 works in the scientific literature.<sup>275</sup>

The SPECIES table is the backbone of FishBase, and has the scientific name as its basic unit. Every bit of information in FishBase is attached directly or indirectly to at least one species and it is mostly through this table that information is accessed. The information in the SPECIES table has been derived from more than 3,500 references.<sup>276</sup>

The SPECIES table presents the valid scientific name and author of a species or subspecies and assigns it to a family, order and class. Where available, a unique English common name is given. Additional information in the SPECIES table relates to maximum age and size, habitat, uses, and general biological remarks. The references used to derive the information are given.

The SPECIES table has links to additional information such as a picture of the fish, a map showing distributional information, higher taxa, synonyms, common names, available life history parameters, all references used, all colleagues who contributed or verified information, etc.<sup>277</sup>

<sup>273</sup> <http://www.fishbase.org/manual/English/fishbaseichthyology00002628.htm>

<sup>274</sup> <http://www.fishbase.org/Photos/PicturesSummary.cfm?ID=10783&what=species> photo by CSIRO

<sup>275</sup> <http://en.wikipedia.org/wiki/FishBase>

<sup>276</sup> [http://www.fishbase.org/manual/English/FishbaseThe\\_Species\\_Table.htm](http://www.fishbase.org/manual/English/FishbaseThe_Species_Table.htm)

<sup>277</sup> [http://www.fishbase.org/manual/English/FishbaseThe\\_Species\\_Table.htm](http://www.fishbase.org/manual/English/FishbaseThe_Species_Table.htm)

FishBase does **not** have:<sup>278</sup>

- Complete checklists (of 296 countries/islands, 70 marine and 140 freshwater checklists were complete in 2000);
- Fish behaviour (only reproductive and trophic behaviour);
- Traditional distribution maps (maps only highlight or mark countries from which a species is reported and plot the occurrence points currently available);
- All references for all species (only list publications that contain suitable information and that have been used so far, currently >20,000);
- Pictures for all fishes (>25,000 pictures for >10,000 species in 2000).

FishBase covers only fish and not other marine forms such as molluscs or crustaceans.<sup>279</sup>

FishBase has been developed at the International Center for Living Aquatic Resources Management (ICLARM) in collaboration with the Food and Agriculture Organization (FAO) of the United Nations and many other partners. FishBase has been funded mainly through sequential grants from the European Commission.<sup>280</sup>

### Function

- Visualisation tools
  - Maps
  - Images
- Taxonomy
  - Identification tools
- User interface
  - Personal and institutional use
  - Raw data and visual presentation

### Why use this tool?

- A biological information source for all fishes

### Who will use this tool?

Similar to an encyclopedia, FishBase contains different things for different people.

For example:<sup>281</sup>

- fisheries managers will dive into the largest existing compilation of population dynamics data;
- teachers and students will find numerous graphs illustrating basic concepts of fish biology;
- taxonomists will enjoy access to the November 2000 update of Eschmeyer's (1998) *Catalog of Fishes* databases;
- conservationists will use the lists of threatened fishes for any given country (Hilton-Taylor 2000);
- policymakers may be interested in a chronological, annotated list of introductions to their country;
- research scientists, as well as funding agencies, will find it useful to gain a quick overview of what is known about a certain species;
- zoologists and physiologists will have the largest existing compilations of fish morphology, metabolism, gill area, brain size, eye pigment, or swimming speed at their fingertips;

---

<sup>278</sup> [http://www.fishbase.org/manual/English/fishbasethings\\_not\\_yet\\_in\\_fishbase.htm](http://www.fishbase.org/manual/English/fishbasethings_not_yet_in_fishbase.htm)

<sup>279</sup> [http://www.fishbase.org/manual/English/fishbasefishbase\\_and\\_groups\\_other\\_than\\_f.htm](http://www.fishbase.org/manual/English/fishbasefishbase_and_groups_other_than_f.htm)

<sup>280</sup> <http://www.fishbase.org/manual/English/index.htm>

<sup>281</sup> <http://www.fishbase.org/manual/English/index.htm>

- ecologists will likewise use data on diet composition, trophic levels, food consumption and predators as inputs for their models;
  - geneticists will find the largest compilation of allele frequencies;
  - the fishing industry will find proximate analyses, as well as processing recommendations for many marine species;
  - anglers will enjoy a listing of all game fishes occurring in a particular country (IGFA 1994);
  - scholars interested in local knowledge will find more than 100,000 common names of fishes together with the language/culture in which they are used and comments on their etymology
- Data capture
    - Curators – specimens, identification
  - Data users
    - Expert
    - Interest groups
    - General public

#### How will the tool be used?

Divers, anglers, aquarists, researchers can create their personal/institutional databases of where and when they have seen, caught, or acquired what fish. Biodiversity managers can create national fish biodiversity databases to keep track of local regulations and uses. Anthropologists can create a database on local knowledge about fish.<sup>282</sup>

- Online and desktop versions
- CD requires windows
- Online data is available as XML
- User input required

#### Where in the data chain could this tool be used?

- Data source
- ALA central
- User's machine

#### When could this tool be used?

- Before data is made available to ALA
- At the time of a user request

#### Availability

- FishBase: <http://www.fishbase.org/>
- FishBase book: [http://www.fishbase.org/Download/FBBook\\_English.zip](http://www.fishbase.org/Download/FBBook_English.zip)
- CD release 2004
- Copyright: International Center for Living Aquatic Resources Management<sup>283</sup>
- Copyright for photos, drawings or arrangements of data remains with the contributors.<sup>284</sup>

#### Comments

- FishBase CD contains a database that could be used as a source of data for the ALA.

<sup>282</sup> <http://www.fishbase.org/manual/English/index.htm>

<sup>283</sup> [http://www.fishbase.org/manual/English/fishbasedisclaimer\\_and\\_copyright.htm](http://www.fishbase.org/manual/English/fishbasedisclaimer_and_copyright.htm)

<sup>284</sup> [http://www.fishbase.org/manual/English/fishbasedisclaimer\\_and\\_copyright.htm](http://www.fishbase.org/manual/English/fishbasedisclaimer_and_copyright.htm)

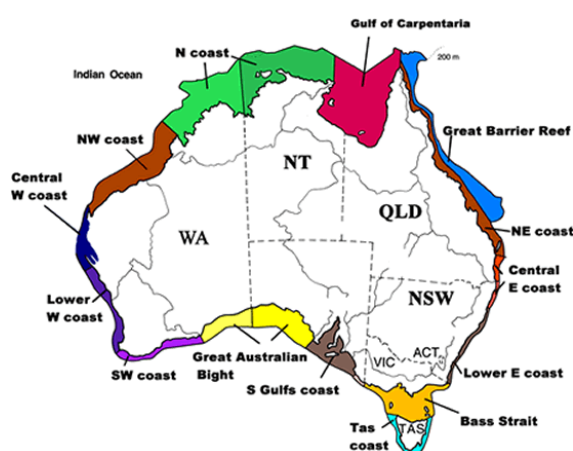
- A link, for example, from an ALA species page could be used to connect that species to an information page on FishBase.
- Uses both common names and scientific names.

## Fishes of Australia

Summary	
Type of tool	Application and database
Function	Species pages, taxonomy keys
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Under development
Time of use	At the time of a user request
Licence	

**Fishes of Australia Online** delivers information on many of Australia's 4500 fish species. Under development.

### Description



*Inshore map for selection of **Fishes by Region**.*<sup>285</sup>

**Fishes of Australia Online** compiles information about the biodiversity and evolutionary relationships of Australian fishes, with the aim of providing information, identification tools, images and other resources for every fish species known from Australian waters.<sup>286</sup>

The vast majority of Australian species have restricted geographical distributions which are determined by their biology, behaviour and historical origins. In most cases, these distributional ranges conform to a number of generalized bioregions around the coastline and to rather regular depth intervals.<sup>287</sup>

**Fish Family Finder** is an interactive aid to the identification of the families of fishes known to occur in the marine and freshwater regions of Australia.

Fishes of Australia online is an initiative of the Australian Ichthyological taxonomy community under the banner of OzFishNet and the Australian Biological Resources Study (ABRS), with support provided by DEW (Department of Environment and Water Resources). Project partners include ABRS, Museum Victoria, CSIRO Marine and Atmospheric Research, the Australian Museum, the Queensland Museum, the Museum and Art Gallery of the Northern

<sup>285</sup> <http://foa.webbooy.net/browse/inshore.cfm>

<sup>286</sup> <http://foa.webbooy.net/about/>

<sup>287</sup> <http://foa.webbooy.net/about/introduction.cfm>

Territory, the Western Australian Museum, the South Australian Museum, Dr Gerald Allen (Conservation International), Rudie Kuitert and Ken Graham (NSW Fisheries).<sup>288</sup>

### Function

- Visualisation tools
  - Maps
  - Images
- Analysis tools
  - Simple
- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

**Fishes of Australia Online** is expected to be used by all people interested in Australian fishes, including students, teachers, nature lovers, amateur scientists, graduate students, lecturers and researchers seeking identification keys, images, phylogenetic tree and other systematic information.<sup>289</sup>

### Who will use this tool?

- Data users
  - Expert
  - Interest groups
  - General public

### How will the tool be used?

- Online application
- User input required

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- At the time of a user request

### Availability

- Fishes of Australia: <http://foa.webboy.net/>
- Identify a fish key: <http://foa.webboy.net/key/>
- Fishes by Region: <http://foa.webboy.net/browse/>
- Example species page: Southern Whiptail *Coelorinchus australis*  
<http://foa.webboy.net/fishes/display.cfm?genus=Coelorinchus&species=australis>
- Contact: <http://foa.webboy.net/contact/>
- Australian Standard Fish Names: <http://www.fishnames.com.au/>

### Comments

---

<sup>288</sup> <http://foa.webboy.net/about/>

<sup>289</sup> <http://foa.webboy.net/about/>

Scientific names and **Australian Standard Fish Names** are used.

## LifeMapper

Summary	
Type of tool	Application
Function	Species distribution modelling
Online / Desktop	Online
Computer infrastructure	
Development status	Experimental. Possibly infant.
Time of use	At the time of a user request
Licence	

**LifeMapper** is building a species diversity map of the world.<sup>290</sup>

### Description

**LifeMapper** uses online geospatial species occurrence data to create distribution maps and, notably, goes one step further to predict where an individual species should exist based on where it is documented to live. LifeMapper does this by combining species occurrence data with global climate, terrain and land cover information, to identify environmental correlates of species ranges.<sup>291</sup>

LifeMapper comprises two primary goals: the construction and maintenance of an extensive predicted species habitat map archive, and the exposure of spatial data and analysis services based on this archive.<sup>292</sup>

The architecture of the LifeMapper project consists of three independent elements. LifeMapper implements the openModeller species niche modeling platform on a cluster of 64 Intel computer nodes with 128 processors and a museum data pipeline to build a global geospatial data archive of predicted species distributions.

The second element of LifeMapper is the workhorse of the project - the data pipeline. The pipeline assembles niche modeling experiments, dispatches them to the openModeller web service, retrieves the results, and catalogs them. This is the element that builds and maintains our archive.

The third element of LifeMapper is the Spatial Data Library (SDL). This is not only an archive of all the input spatial data used in creating the habitat maps, but also a catalog of the resulting niche model maps. Data in the SDL is publicly available via REST web services for the metadata and OGC services for the spatial data. The website provides a mechanism for browsing the archive and exploring environmental data, species occurrence points, and niche model maps while web services built on the archive are targeted at researchers who would like to programmatically query, analyse, and download the data produced.<sup>293</sup>

These elements make up the LifeMapper project. They can operate in tandem or independently: each element could be replaced by a comparable service or application for a similar output, or incorporated into a new application with unique objectives.

---

<sup>290</sup> <http://en.wikipedia.org/wiki/Lifemapper>

<sup>291</sup> <http://www.lifemapper.org/index.shtml>

<sup>292</sup> <http://www.lifemapper.org/technology.shtml>

<sup>293</sup> <http://www.lifemapper.org/technology.shtml>



## Web Services<sup>294</sup>

LifeMapper will connect existing computational efforts in environmental biology by providing open access network interfaces for web service integration with client applications and other biodiversity web servers.

LifeMapper is funded by US NSF and has partners: [CRU](#), [GBIF](#), [IPCC](#), [Morphbank](#), [openModeller](#)<sup>295</sup>

### Function

- Visualisation tools
  - Maps
- Analysis tools
  - distribution
- Taxonomy
  - Identification tools, keys
- User interface
  - Personal and institutional use
  - Raw data and visual presentation

### Why use this tool?

- For species distribution maps

### Who will use this tool?

- Data users
  - Expert
  - Interest groups

### How will the tool be used?

LifeMapper exposes web services that return xml documents or spatial data from a simple URL query. These services can be used by any application, one of which is the LifeMapper website. The LifeMapper project will initially offer some basic web services to the public.

- list environmental scenarios and their layers
- query for environmental scenarios by keywords or spatial extent
- query for layers by keywords or spatial extent
  
- list Occurrence Sets and their points
- query for Occurrence Sets by taxonomic name or spatial extent
  
- list models
- query for models by taxonomic name or algorithm
  
- list predicted habitat maps
- query for predicted habitat maps by taxonomic name, algorithm, scenario, or spatial extent
  
- render or download spatial data for any of the above using OGC standards WFS or WCS
  
- request modeling or projection services using user-provided or LifeMapper cached data

---

<sup>294</sup> <http://www.lifemapper.org/webServices.shtml>

<sup>295</sup> <http://www.lifemapper.org/about.shtml>

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- At the time of a user request

**Availability**

- LifeMapper: <http://www.lifemapper.org/>
- Contact: Jim Beach – Director – beach @ku.edu
- Experimental

**Comments**

- From the LifeMapper website there does not appear to be much life happening at LifeMapper.
- See also: Chapman, A.D. (2004m). LifeMapper. Comments and Ideas. Appendix K to */Sistema de Informação Distribuído para Coleções Biológicas: A Integração do Species Analyst e SinBiota. FAPESP/Biota process no. 2001/02175-5 March 2003 – March 2004./* Campinas, Brazil: CRIA 11 pp. [http://smlink.cria.org.br/docs/appendix\\_k.pdf](http://smlink.cria.org.br/docs/appendix_k.pdf)

## Lucid keys

Summary	
Type of tool	Taxonomy tool
Function	Identification and diagnostics
Online / Desktop	Online
Computer infrastructure	Web based and Java based, platform independent
Development status	Operational. Version 3
Time of use	As a post process, when data is with user
Licence	

**Lucid** tools are powerful and highly flexible knowledge management software applications designed to help users with identification or diagnostic tasks.<sup>296</sup>

### Description

Lucid software is a special type of expert system, specifically designed for identification and diagnostic purposes, which enables expert knowledge to be “cloned” and distributed to a wide audience via CD or the Internet. The large number of functions incorporated in the software and the ability to include multi-media makes the creation and use of identification and diagnostic keys easy, effective and enjoyable.

Lucid identification keys are currently being used by a wide range of end-users, from primary, high school and university students to taxonomists, quarantine identifiers, biodiversity scientists and conservation managers.<sup>297</sup>

**Lucid3** users who wish to identify an entity in the group for which the key has been built can "describe" their specimen to the key by choosing features that match the specimen. The key progressively eliminates entities that do not match the chosen features until only one or a few remain - multimedia fact sheets and images provide further information for making an identification.<sup>298</sup>

**Lucid Phoenix** is a computer based *dichotomous* or *pathway* key Builder and Player that enables traditional paper based identification keys to be published on the Internet or CD. Phoenix keys are interactive, can be enhanced with multimedia, and delivered across the Internet seamlessly.<sup>299</sup>

The Lucid software development team is located at the Centre for Biological Information Technology (CBIT) at the University of Queensland, Australia.<sup>300</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use

---

<sup>296</sup> <http://www.lucidcentral.org/about/>

<sup>297</sup> <http://www.lucidcentral.org/about/>

<sup>298</sup> <http://www.lucidcentral.org/lucid3/>

<sup>299</sup> <http://www.lucidcentral.org/phoenix/default.htm>

<sup>300</sup> <http://www.lucidcentral.org/about/us.htm>

### **Why use this tool?**

- Identification or diagnostic

### **Who will use this tool?**

- Data users
  - Expert
  - Interest groups
  - General public

### **How will the tool be used?**

Lucid tools are available both as online keys and distributed as a desktop application on CD. Of most interest to the ALA would be the online keys.

The Lucid keys can be used online through one of two Players – the Lucid3 On-line Player and the Lucid3 Applet.

The keys can be hosted on any internet domain.

#### **Lucid3 On-line Player**

- Requires a web server to deliver content, eg Linux/Apache or Windows Server/IIS. Uses Tomcat and Java based technology.
- Browser based, platform independent
- Will work on PDA and mobile phones
- Uses Standard for Descriptive Data (SDD) data format
- Currently under development, in pre-release phase

#### **Lucid3 Applet**

- Java applet
- Browser based, platform independent
- User input is required

#### **Lucid Phoenix**

- Phoenix keys can be directly embedded within web pages
- Cross platform, operates on Windows, Mac, Unix
- Java required

### **Where in the data chain could this tool be used?**

- User's machine

### **When could this tool be used?**

- As a post process, after data is with the user

### **Availability**

- Lucid: <http://www.lucidcentral.org/>
- Lucid3: <http://www.lucidcentral.org/lucid3/>
- Lucid Phoenix: <http://www.lucidcentral.org/phoenix/>
- CBIT Centre for Biological Information Technology: <http://www.cbit.uq.edu.au/>
- Support: <http://www.lucidcentral.org/support/>
- No licence fees are required when deploying a key, non-commercially, to a free internet site.
- Commercial distribution attracts a fee. Contact [sales@lucidcentral.org](mailto:sales@lucidcentral.org)
- Free key hosting service at Lucid

Phoenix Pupal Key to Genera of White Flies - Created using Lucid Phoenix - <http://www.lucidcentral.com/phoenix/keys/Pupal%20Key%20to%20Genera%20of%20White%20Flies.htm>

File Edit View Favorites Tools Help

Back Forward Stop Home Search Favorites Media Print Mail

Address <http://www.lucidcentral.com/phoenix/keys/Pupal%20Key%20to%20Genera%20of%20White%20Flies.htm> Go Links

Questions

[Two pairs of submedial, 2-segmented siphon like tubes widen at apices on meso- and metathorax](#)

[Many siphons on dorsum](#)

History

- Thoracic legs without claw (Aleyrodinae)
  - Skipped question
    - Long siphons (wax tubes) or segmented siphon like tubes with expanded apex present
      - Pending question
    - Siphons or siphon-like tubes absent; with spines or setae (simple or otherwise)
      - Pending question

Entities Remaining: 44 (80%)

- Siphoninus phillyreae (Haliday)
- Aleurothrixus antidesmae Takahashi
- Trialeurodes
- Aleurotrachelus
- Aleyrodes
- Aleurotuberculatus
- Acaudaleyrodes
- Aleuroparadoxus
- Aleurocanthus
- Aleuroglandulus
- Aleurotuba
- Bemisia
- Corbettia
- Cratichneumon

Entities Discarded: 11 (20%)

- Metaleurodicus
- Lecanoideus
- n.genus
- Dialeurodicus
- Bakerius
- Paraleyrodes
- Octaleurodicus
- Aleurodicus
- Bondaria
- Ceraleurodicus
- Aleuronudus

Applet com.cbit.phoenix.player.PhoenixApplet started

Trusted sites

Lucid Phoenix player within a browser.<sup>301</sup>

## Comments

Lucid Phoenix is quite a different system from Lucid3. It's never really been properly taken up, partly because many people have the misapprehension that dichotomous keys are primitive. In fact they have some striking advantages over matrix keys, and the idea of Phoenix is to draw out the strengths while minimizing the weaknesses.<sup>302</sup>

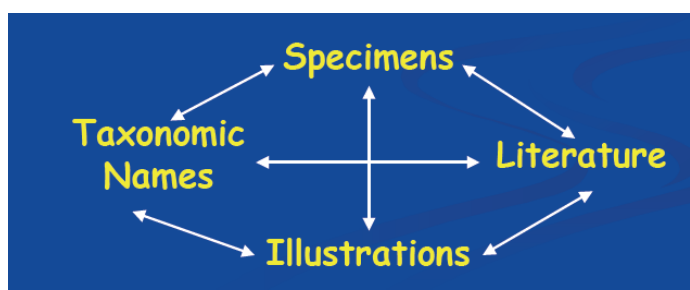
<sup>301</sup> <http://www.lucidcentral.org/phoenix/screenshots.htm>

<sup>302</sup> Kevin Thiele, Western Australia Department of Environment and Conservation, February 2008

## Mandala

Summary	
Type of tool	Application
Function	Data management, georeferencing
Online / Desktop	Desktop
Computer infrastructure	Mac or Windows, FileMaker Pro 7/8.x
Development status	Active development 1995-2007. Version 7.
Time of use	Data preparation, post process after data is with user
Licence	Non-commercial research and internal business use

**Mandala** is a suite of 23 interrelated database tables (see [Overview](#) and [Model](#)), which uses the cross-platform database application, FileMaker Pro. Mandala supports four major realms of data acquisition and management for systematics and biodiversity studies: specimens, literature, taxonomic names, and illustrations.<sup>303</sup>



*Mandala data acquisition and management*<sup>304</sup>

### Description

Mandala can be used to:<sup>305</sup>

- track museum loans, specimens, bulk samples & their subsamples;
- detail the complex history of a taxonomic name;
- export a multitude of data for phenological plots, specimens examined lists, and distribution maps;
- provide direct URLs to GenBank records;
- catalogue images and illustrations;
- store and link literature to appropriate taxonomic names, specimens, illustrations, and collecting localities; and
- record numerous details about individual specimens including a literal transcription of the label(s), enhanced locality and collecting event data as interpreted from the label(s), taxonomic determination history, physical condition of the specimen and how it has been preserved, ecological/biological associations with other collected specimens or other taxa associated but not collected, type designation, and other relevant information.

Mandala is a research tool and data manager. New data may be constantly added, enhanced, and updated. Ongoing structural changes from user requests facilitate data input and enhance ease of output. Data in Mandala may be in various stages of verification and completeness.

<sup>303</sup> <http://www.inhs.uiuc.edu/research/mandala/about.html>

<sup>304</sup> <http://www.inhs.uiuc.edu/research/mandala/Overview.pdf>

<sup>305</sup> <http://www.inhs.uiuc.edu/research/mandala/about.html>

Specimens in Mandala have unique numbers that are linked to taxonomic names, localities and collecting events, illustrations, literature, associated specimens, other associated organisms, ecological and behavioural observations via a controlled vocabulary, biogeographic regions, loan and deposit information from collections, determinations, and a wealth of other characteristics about the physical specimen itself, including sex, mounted state, dissections, GENBANK Accession Numbers and URLs, developmental stage collected and in collection, and pupation and emergence dates.

The taxonomic names file (TAXA) allows the user to document all facets of name use, including references verifying that use. The history of a taxonomic name can be extremely complex, requiring the attention of a professional systematist, yet a taxonomic record does not need to be complete to be functional e.g., attaching a working name to a specimen. The construction of Mandala allows not only the tracking of valid and invalid taxonomic names but also of working, manuscript, and in press names. The relational framework of Mandala allows the automatic display of homonyms as well as a full synonymic list of names linked to any valid name, which may be exported for nomenclatural catalogues.

Created as part of a systematics project on stiletto flies (Insecta: Diptera: Therevidae), the development of Mandala was funded by the National Science Foundation's Partnerships for Enhancing Expertise in Taxonomy programs, the NSF-sponsored Fiji Bioinventory of Arthropods and the Schlinger Foundation. It was developed by Gail Kampmeier.<sup>306</sup>

#### **Function**

- Data management
  - Data validating – taxonomy, geography, duplication
  - Georeferencing – applying latitude and longitude
- Taxonomy
- Provider interaction
  - Data preparation
- User interface
  - Personal and institutional
  - Raw data and visual presentation

#### **Why use this tool?**

- Georeferencing
- Data management
- Webserver options<sup>307</sup>

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
- Data users
  - Expert

---

<sup>306</sup> <http://www.inhs.uiuc.edu/research/mandala/about.html>

<sup>307</sup> <http://www.inhs.uiuc.edu/research/mandala/Overview.pdf>

- Interest groups

**How will the tool be used?**

- Desktop application
- Windows or Mac, Filemaker Pro 7/8.x
- Access online Gazetteers and Taxonomic Names Servers
- User input required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- Before data is made available to ALA
- As a post process, after data is with the user

**Availability**

- Mandala: <http://www.inhs.uiuc.edu/research/mandala/>
- Overview: <http://www.inhs.uiuc.edu/research/mandala/Overview.pdf>
- Contact: Gail Kampmeier [gkamp@uiuc.edu](mailto:gkamp@uiuc.edu)
- Version 7
- Cost: free
- Licence: <http://www.inhs.uiuc.edu/~gkamp/downloads/MandalaLicense.pdf>
- Restrictions: Non-commercial research and internal business use.

**Comments**

Mandala is a collections front-end to FileMaker Pro. It has been developed by and used strongly by the fly research community.

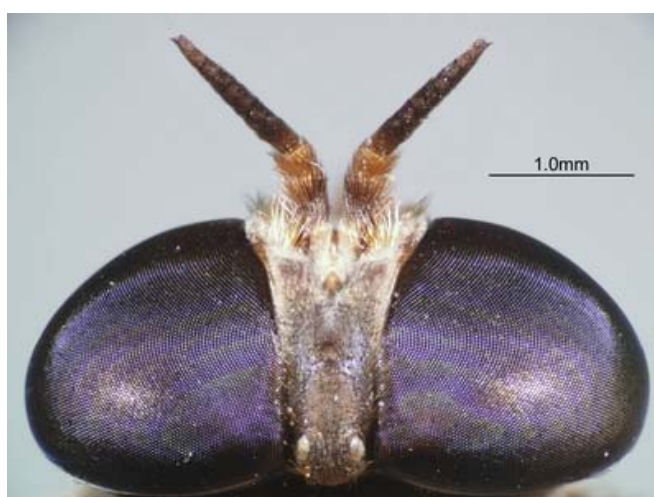


## Morphbank

Summary	
Type of tool	Database of images
Function	Taxonomy
Online / Desktop	Online
Computer infrastructure	Browser with pop-ups, cookies, JavaScript, Java, pdf
Development status	Active development. Version 2.7, November 2007
Time of use	At the time of a user request
Licence	Creative commons

**Morphbank** is a continuously growing database of images that scientists use for international collaboration, research and education.<sup>308</sup>

### Description



*Exeretonevra angustifrons*, Bago Forest, Batlow, NSW. Christine Lambkin<sup>309</sup>

Images deposited in Morphbank document a wide variety of research including: specimen-based research in comparative anatomy, morphological phylogenetics, taxonomy and related fields focused on increasing our knowledge about biodiversity.

Overall, Morphbank facilitates research efforts by making it possible to store, discuss and share detailed images of specimens from all over the world. Currently, Morphbank holds more than 63,000 public images of more than 4500 different species. Many more images are in the system that are presently not public, but are held private until the contributing scientists are ready to release the images. [They use] these images to document characters and features to describe organisms, make observations about the organisms, or make a comment about the images themselves. The set of morphological features unique to a specimen can be used to identify future specimens. New species are discovered and described by looking at images in Morphbank and extinct or very rare specimens are preserved for current and future generations to study.

Morphbank was established in 1998 by a Swedish-Spanish-American group of entomologists and is currently housed at the School of Computational Science (SCS) at Florida State University. The

<sup>308</sup> <http://www.morphbank.net/About/AboutMb/>

<sup>309</sup> <http://www.morphbank.net/Show/?pop=Yes&id=196354>

project has grown immensely since its beginnings and presently includes a team of 15 biologists, computer scientists and information scientists who are working on developing the software. Morphbank is dedicated to using open-source software and is a Fair Use Web Site.

The project receives its main funding from the Biological Databases and Informatics program of the National Science Foundation.<sup>310</sup>

### **Function**

- Data cleaning and manipulation
  - Data validating – taxonomy
- Visualisation tools
  - Images
- Taxonomy
  - Identification
- Provider interaction
  - Data preparation
- User interface
  - Personal and institutional
  - Visual presentation

### **Why use this tool?**

Morphbank is an excellent source of information for identification of biological specimens.

In taxonomy, descriptions of new species or other nomenclatural acts can be documented by images and image comment tools in Morphbank. Biologists who use Morphbank can mark an image, make comments about images that are seen by other researchers often continents away.

Morphbank provides stable Web links to image collections for use in online and traditional print publications. Large biological inventory projects can use Morphbank to share image information about morphological types and allow experts to do remote identification.<sup>311</sup>

### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - Interest groups
  - General public
- ALA infrastructure

### **How will the tool be used?**

---

<sup>310</sup> <http://www.morphbank.net/About/AboutMb/>

<sup>311</sup> <http://www.morphbank.net/About/AboutMb/>

Morphbank users who wish to access, add and modify data will be required to login into the system.<sup>312</sup> Members are assigned appropriate taxonomic privileges. Without login, only published (public) information that can be browsed or searched by the casual user (specimen, image, publication, locality, view, or collection data) can be seen.<sup>313</sup>

- Online searchable database of images
- Upload: web interface, ftp, CD and Excel spreadsheet<sup>314</sup>
- Browser required with pop-ups and cookies, Java and JavaScript, pdf, media player<sup>315</sup>
- User input is required
- Batch mode for submission

#### **Where in the data chain could this tool be used?**

- Data source
- ALA central
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- While data is stored with ALA
- At the time of a user request

#### **Availability**

- Morphbank: <http://www.morphbank.net/>
- User manual: <http://www.morphbank.net/About/Manual/>
- Documentation: <http://www.morphbank.net/Help/Documents/>
- ENBI - Digital Imaging of Biological Type Specimens - A Manual of Best Practice: [http://circa.gbif.net/Public/irc/enbi/comm/library?l=/enbi\\_reports/biological\\_specimens&vm=detailed&sb=Title](http://circa.gbif.net/Public/irc/enbi/comm/library?l=/enbi_reports/biological_specimens&vm=detailed&sb=Title)
- Version 2.7, November 2007
- Cost: free
- Copyright: <http://www.morphbank.net/About/Copyright/>
- Licence: [Attribution-Noncommercial-Share Alike 3.0 Creative Commons by-nc-sa](http://creativecommons.org/licenses/by-nc-sa/3.0/) for every image:<sup>316</sup>

#### **Comments**

---

<sup>312</sup> <http://www.morphbank.net/About/Manual/userPrivileges.php>

<sup>313</sup> <http://www.morphbank.net/About/Manual/userPrivileges.php>

<sup>314</sup> <http://www.morphbank.net/About/HowToContribute/>

<sup>315</sup> <http://www.morphbank.net/About/Manual/systemRequire.php>

<sup>316</sup> <http://www.morphbank.net/Show/?id=199848>

## NBN Species Dictionary

Summary	
Type of tool	Example of a searchable dictionary
Function	Common or scientific name search, taxonomy
Online / Desktop	Online
Computer infrastructure	Browser with JavaScript, pdf.
Development status	Active in development. Updated July 2007
Time of use	At time of user request
Licence	Personal and non-commercial use. Over 18 years, or consent of parent.

The **NBN Species Dictionary** aims to provide a standard reference for names of organisms found in the United Kingdom.<sup>317</sup> This is an example of a taxon dictionary.

### Description

The British National Biodiversity Network Species Dictionary is made up from a collection of over 210 separate checklists containing more than 196,000 versions of taxonomic names and other associated information. You can either search individual lists or find organisms by their common or scientific name. A basic classification of life is also provided as a navigational tool.<sup>318</sup>

To share information about animals and plants we must be able to translate between the names used in different sources of information and to relate different scientific names used for the same species. The Species Dictionary aims to do just this by bringing all of the names together in one place, and providing a means for translating between them. Although still under development, it is aimed that the Project will become the focal point for a series of comprehensive and authoritative checklists covering the entire flora and fauna of the British Isles. It will thus ultimately provide a complete and definitive list of the scientific names of all British organisms.

The most important feature of the Species Dictionary is that it is not a single checklist of taxa but a mechanism for storing many checklists and versions of checklists, together with the means for translating from one to another. This is what makes it very different from virtually all other taxonomic database projects. The importance of this is that biological records should be stored with their original determinations to which re-determinations may be added. It should be possible for the software to retrieve taxa using correspondences of names in alternative checklists.

Over 210 different lists (of which 154 are searchable through this website), covering various groups or regions, are presently incorporated and this number will continue to grow over time. This information is compiled from a wide range of sources and as such the Species Dictionary relies on countless hours of work by those who have compiled the lists. In recognition of this, the authors and compilers of the various datasets are fully acknowledged throughout. The National Biodiversity Network (NBN) Species Dictionary currently contains over 196,000 names.<sup>319</sup>

Other points of interest:<sup>320</sup>

---

<sup>317</sup> <http://nbn.nhm.ac.uk/nhm/>

<sup>318</sup> <http://nbn.nhm.ac.uk/nhm/>

<sup>319</sup> <http://nbn.nhm.ac.uk/nhm/info.shtml#project>

<sup>320</sup> <http://nbn.nhm.ac.uk/nhm/info.shtml#project>

- The Species Dictionary website allows users to search for names of organisms by their common and scientific names, and through an hierarchical classification. You can also search for individual lists and their contents.
- The database serves as an archive by presenting various versions and editions of a number of important checklists, thereby allowing the user to investigate how the names, or threat statuses, of taxa within particular groups of organisms have changed through time.
- The database also holds information concerning the legal designations and conservation status of protected species.
- The Species Dictionary is incorporated within a number of other NBN products. In addition to Recorder, it forms part of the [NBN Gateway](#) and [NBN Index](#).

This is a project of the National Biodiversity Network and is hosted and managed by the Natural History Museum, London.<sup>321</sup>

### Function

- Data cleaning
  - Data validating – taxonomy
- Taxonomy
- User interface
  - Personal and institutional use
  - Raw data and visual presentation

### Why use this tool?

- A means for translating between common names and valid scientific names
- A method of accessing a master names list

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - Interest groups
  - General public
- ALA infrastructure
- No special skills required – accepts both common names and scientific names

### How will the tool be used?

- Online searchable dictionary
- A navigable hierarchical classification
- Contains a large and increasing number of checklists<sup>322</sup>
- Browser with JavaScript, pdf
- User input is required

<sup>321</sup> <http://nbn.nhm.ac.uk/nhm/>

<sup>322</sup> [http://nbn.nhm.ac.uk/nhm/summary\\_table.pdf](http://nbn.nhm.ac.uk/nhm/summary_table.pdf)

**Where in the data chain could this tool be used?**

- ALA central – as delivery
- User's machine – as access

**When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request

**Availability**

- NBN Species Dictionary: <http://nbn.nhm.ac.uk/nhm/>
- Hierarchical classification: <http://nbn.nhm.ac.uk/nhm/kingdoms.shtml>
- Browse species by common names:  
<http://www.searchnbn.net/directory/browseTGLevel1.jsp?consFilter=0&dataFilter=1>
- NBN – National Biodiversity Network: <http://www.nbn.org.uk/>
- Summary of checklists: [http://nbn.nhm.ac.uk/nhm/summary\\_table.pdf](http://nbn.nhm.ac.uk/nhm/summary_table.pdf)
- Cost: free
- Under active development. Updated July 2007.
- Restrictions: Personal and non-commercial use. Over 18 years, or consent of parent.
- Terms of use <http://www.nhm.ac.uk/about-us/website-help/terms-of-use/index.html>

**Comments**

The NBN Species Dictionary, although specific to Britain, has been included as an example of a searchable taxon dictionary. It allows for both common names and scientific names, and variations of these names over place and time.

Feedback opportunities are given to alert to taxonomic inconsistencies or mistakes.

## PaDIL – Pest and Disease Image Library

Summary	
Type of tool	Database of high resolution images
Function	Viewing pest images, taxonomy of pests
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Active
Time of use	At time of user request
Licence	Research, education, and personal use.

High-quality colour diagnostic images and information on pests and diseases. **PaDIL** helps protect against invasive threats to Australia's plant health.<sup>323</sup>

### Description

**PaDIL – Pest and Disease Image Library**, combines leading edge technology in light microscopy, digital imaging and image manipulation. Utilising an innovative, non-molecular mechanism, it provides 'virtual specimens' of a type only previously possible using low-power scanning-electron microscopy.<sup>324</sup>



Red fire ant *Solenopsis invicta*.<sup>325</sup>

- A series of images is captured, each providing an in-focus image of part of the photographed organism.
- Each series of images is then combined to create a single new image, built from only the in-focus pixels from each individual image.
- Specimens are shown in true-life colour.
- Destructive preparation of specimens is not required.
- Diagnostic images show the user the best orientation to view the specimen.
- Results allow a user with a microscope to focus on different parts of the viewed organism, simultaneously comparing what they see against a single PaDIL image.

<sup>323</sup> <http://www.padil.gov.au/>

<sup>324</sup> <http://www.padil.gov.au/aboutSolution.aspx>

<sup>325</sup> <http://www.padil.gov.au/viewPestDiagnosticImages.aspx?id=93>

- All images are copyright free for non-commercial purposes.

PaDIL is a Commonwealth Government initiative, developed and built by Museum Victoria's Online Publishing Team, with support provided by DAFF (Department of Agriculture, Fisheries and Forestry) and PHA (Plant Health Australia), a non-profit public company.<sup>326</sup>

### Function

- Visualisation tools
  - Images
- Taxonomy
  - Identification tools
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

Quarantine and plant health are critical to the economic, social and environmental well-being of Australia.<sup>327</sup>

- Australia's existing taxonomic expertise is declining and aging.
- From a plant health and quarantine perspective, this brings an increasing risk that incursions will go unrecognised or be diagnosed too late to be contained.

PaDIL helps alleviate the loss of real-world expertise through the generation of 'virtual' expertise.<sup>328</sup>

- Virtual collections of high quality digital images, illustrated diagnostic keys and on-line tutorials overcome the expense of maintaining centralised technical reference collections.
- Illustrations are focussed on key diagnostic features.
- Diagnostic material is available to a much larger group of users.
- Taxonomic expertise is unnecessary, particularly where diagnosis is aimed at a limited group of target species (i.e. pest species on a quarantine target list).

### Who will use this tool?

- Data users
  - Expert
  - Interest groups
  - General public
- Special skills not required

### How will the tool be used?

- Online database of images
- Web Browser
- Is it used on the desktop or online?
- User input required
- Not a batch job

### Where in the data chain could this tool be used?

- User's machine

<sup>326</sup> <http://www.padil.gov.au/aboutOverview.aspx>

<sup>327</sup> <http://www.padil.gov.au/aboutBackground.aspx>

<sup>328</sup> <http://www.padil.gov.au/aboutSolution.aspx>



**When could this tool be used?**

- At the time of a user request

**Availability**

- PaDIL: <http://www.padil.gov.au>
- Help: <http://www.padil.gov.au/helpMain.aspx>
- Cost: free
- Restrictions: Saving and printing acceptable for research, education, and personal use. Material is not available to be sold or re-sold without permission.
- Copyright: Owned by authors

**Comments**

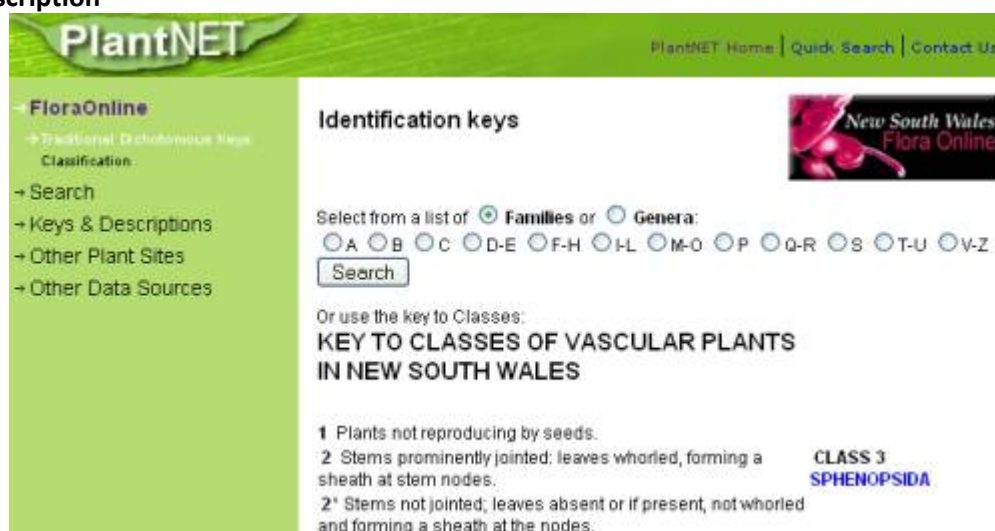
Impressive high quality images with the ability to pan and zoom.

## PlantNET

Summary	
Type of tool	Information system
Function	Plant taxonomy
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Operational since 1999. Now version 2.0 March 2004
Time of use	At time of user request
Licence	Personal, in-house and non-commercial use

**PlantNET** is an information system that can be searched for information on New South Wales plant names, their distribution and conservation status. Identification tools (keys), descriptions, images, illustrations and other information are provided.<sup>329</sup>

### Description



The screenshot shows the PlantNET website interface. At the top, there is a green header with the 'PlantNET' logo and navigation links for 'PlantNET Home', 'Quick Search', and 'Contact Us'. On the left side, there is a green sidebar menu with the following items: 'FloraOnline', '→ Traditional Dichotomous Keys', 'Classification', '→ Search', '→ Keys & Descriptions', '→ Other Plant Sites', and '→ Other Data Sources'. The main content area is titled 'Identification keys' and features a search interface. It includes a radio button selection for 'Families' (selected) or 'Genera', followed by a list of letters from A to Z with radio buttons. Below this is a 'Search' button. Underneath, it says 'Or use the key to Classes:' followed by the title 'KEY TO CLASSES OF VASCULAR PLANTS IN NEW SOUTH WALES'. A list of characteristics is shown, with 'CLASS 3 SPHENOPSIDA' highlighted in blue.

PlantNET consists of several modules that provide information of the plants of New South Wales.<sup>330</sup>

- Search by name: accepted and unaccepted scientific name, common (vernacular) names and groups
- Search by location: using botanical regions of the National Herbarium of New South Wales, or advanced searching based on latitude/longitude coordinates
- Use identification keys for unknown plants of selected or defined areas of New South Wales
- View descriptions, notes on distribution, other information, images

These modules include:

- NSWplants - search for a plant by name, location, or using various characteristics
- HerbLink - search for images of type specimens held at the National Herbarium of New South Wales
- WeedAlert - lists new distributional records of introduced plants to New South Wales in the last month, last three months or last 12 months

<sup>329</sup> <http://plantnet.rbgsyd.nsw.gov.au/about.html>

<sup>330</sup> <http://plantnet.rbgsyd.nsw.gov.au/howto.html>

- FloraOnline - Advanced area and botanical searches, pre-determined and user defined interactive keys to unknown plants, descriptions, images, detailed distribution maps based on herbarium records, and other facts
- Eucalink - providing specialised information on eucalypts (Angophora, Corymbia, and Eucalyptus); including keys, descriptions, images & maps
- WattleWeb - providing specialised information on wattles of New South Wales; including keys, descriptions, images & maps

The data presented in *PlantNET* is generally the most current and authoritative scientific information available on the nomenclature, taxonomy and distribution of the plants of New South Wales or other areas as represented by data on sites hosted by *PlantNET*. The National Herbarium of New South Wales (NSW) updates name and specimen records daily, likewise images and descriptive text are maintained and updated regularly.<sup>331</sup>

#### **Function**

- Visualisation tools
  - Maps
  - Images
- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Visual presentation

#### **Why use this tool?**

- NSW plant information

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data users
  - Expert
  - Interest groups
  - General public
- Some specialist skills may be required

#### **How will the tool be used?**

- Online tool
- User input required

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- At the time of a user request

#### **Availability**

---

<sup>331</sup> <http://plantnet.rbgsyd.nsw.gov.au/limitofdata.html>

- PlantNET: <http://plantnet.rbgsyd.nsw.gov.au/>
- NSW FloraOnline: <http://plantnet.rbgsyd.nsw.gov.au/floraonline.htm>
- Search: <http://plantnet.rbgsyd.nsw.gov.au/search/index.html>
- FloraSearch NSW plants: <http://plantnet.rbgsyd.nsw.gov.au/search/florasearch.htm>
- HerbLINK images of type specimens: <http://plantnet.rbgsyd.nsw.gov.au/search/herblink.htm>
- WeedAlert: <http://plantnet.rbgsyd.nsw.gov.au/WeedAlert/>
- Keys and descriptions: [http://plantnet.rbgsyd.nsw.gov.au/keys\\_desc.htm](http://plantnet.rbgsyd.nsw.gov.au/keys_desc.htm)
- Operational since 1999. Now version 2.0 March 2004
- Copyright: personal, in-house and non-commercial use

## Comments

## **Name resolution**

## AFD – Australian Faunal Directory

Summary	
Type of tool	Searchable database
Function	Taxonomy, curatorial
Online / Desktop	Online database
Computer infrastructure	Web browser
Development status	Operational
Time of use	Data preparation, data processing
Licence	Free

**Australian Faunal Directory (AFD)** is a public enquiry database and will serve as a source of taxonomic and biological information on all animal species known to occur in Australia.<sup>332</sup> This is a project of Australian Biological Resources Study (ABRS).

### Description

#### Main Checklist

Central to the *Australian Faunal Directory* is a checklist which aims to list to family all animal groups known to occur in Australia and its dependant territories. Linked to this checklist at various taxonomic levels are databases ranging from simple checklists of names through to comprehensive catalogue style reports. Links to other sources of information, such as the *Fauna of Australia*, identification keys, endangered species, are under development.

#### Content

During compilation of the *Australian Faunal Directory*, the level of information available will vary from one animal group to another. For some groups, this may consist initially of only a list of scientific names to species. For others it will include the data compiled for the *Zoological Catalogue of Australia* database and further information on biology, ecology and distribution. Where species lists are not yet available, the *Australian Faunal Directory* will aim to list at least all families known to occur in Australia.

#### Status

Valid names are now available for an estimated 55% of described species. The groups available in the *Australian Faunal Directory* are listed on the Groups page. Data from both the *Zoological Catalogue of Australia* and the *Fauna of Australia* are included in the *Australian Faunal Directory*.

#### Source of Data

The central checklist to family was compiled from various sources and includes data derived from the *Zoological Catalogue of Australia* database, ABRS publications and advice from specialists. Due to the uncertain and changing status of many taxa, this checklist should be treated only as a guide and as a means to present the data. The checklist will be updated as new information is received. Published sections of the *Catalogue* database have been peer reviewed but users are advised that revised sections may not have been reviewed.<sup>333</sup>

### Function

- Taxonomy
  - Identification tools, keys

<sup>332</sup> <http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/afd/index.html#Platform>

<sup>333</sup> <http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/afd/index.html#Platform>

- User interface
  - Personal and institutional use

#### **Why use this tool?**

- An authoritative source of scientific names for taxonomy

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data users
  - Expert
  - Interest groups
- ALA infrastructure

#### **How will the tool be used?**

- Data can be downloaded as a RTF or HTML table for importing into a spreadsheet.
- Online tool
- User input required

#### **Where in the data chain could this tool be used?**

- Data source
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request
- As a post process, after data is with the user

#### **Availability**

- Australian Faunal Directory:  
<http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/afd/>
- ABRS Fauna Online: <http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/>
- Checklist:  
<http://www.environment.gov.au/cgi-bin/abrs/fauna/tree.pl?pstrVol=TOC&pintMode=65>
- Copyright © Commonwealth of Australia, Department of Environment and Water Resources:  
<http://www.environment.gov.au/about/copyright.html>

#### **Comments**

## APNI – Australian Plant Names Index

Summary	
Type of tool	Searchable database
Function	Taxonomy, curatorial
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Active since 1991
Time of use	Data preparation, data processing
Licence	Non-commercial use

**APNI - Australian Plant Name Index** is a list of all published names of Australian vascular plants and their bibliographic and typification details.<sup>334</sup>

### Description

APNI is the standard dataset for plant names in Australia, containing more than 80,000 names.<sup>335</sup>

The APNI database is a tool for the botanical community that deals with plant names, not only those in current use, but also those names used in the past – synonyms no longer in use for a variety of technical nomenclatural reasons. It includes information on:<sup>336</sup>

- names in current use;
- plant name changes;
- previous names for a renamed plant;
- the journal or place where the name was formally published;
- the author of the plant name;
- the place in Australia where the Type of the name was collected;
- the Herbarium where the Type Specimen is lodged;
- additional references;
- relevant comments and notes on the naming process, and
- links to other information such as plant distributions, descriptions and images.

APNI does not offer an 'opinion' on which of the names are currently accepted.<sup>337</sup>

It is recognised by Australian herbaria as the prime reference for Australian plants, not only for taxonomy, but also for fields such as ecology and horticulture; it forms the basis for legislation relating to the natural environment. Maintenance of APNI is strongly supported by the Council of Heads of Australian Herbaria (CHAH), the peak body representing Australian herbaria, and the botanical community as a whole. APNI is an integral component of **Australia's Virtual Herbarium (AVH)**.<sup>338</sup>

---

<sup>334</sup> <http://www.anbg.gov.au/cpbr/databases/apni.html>

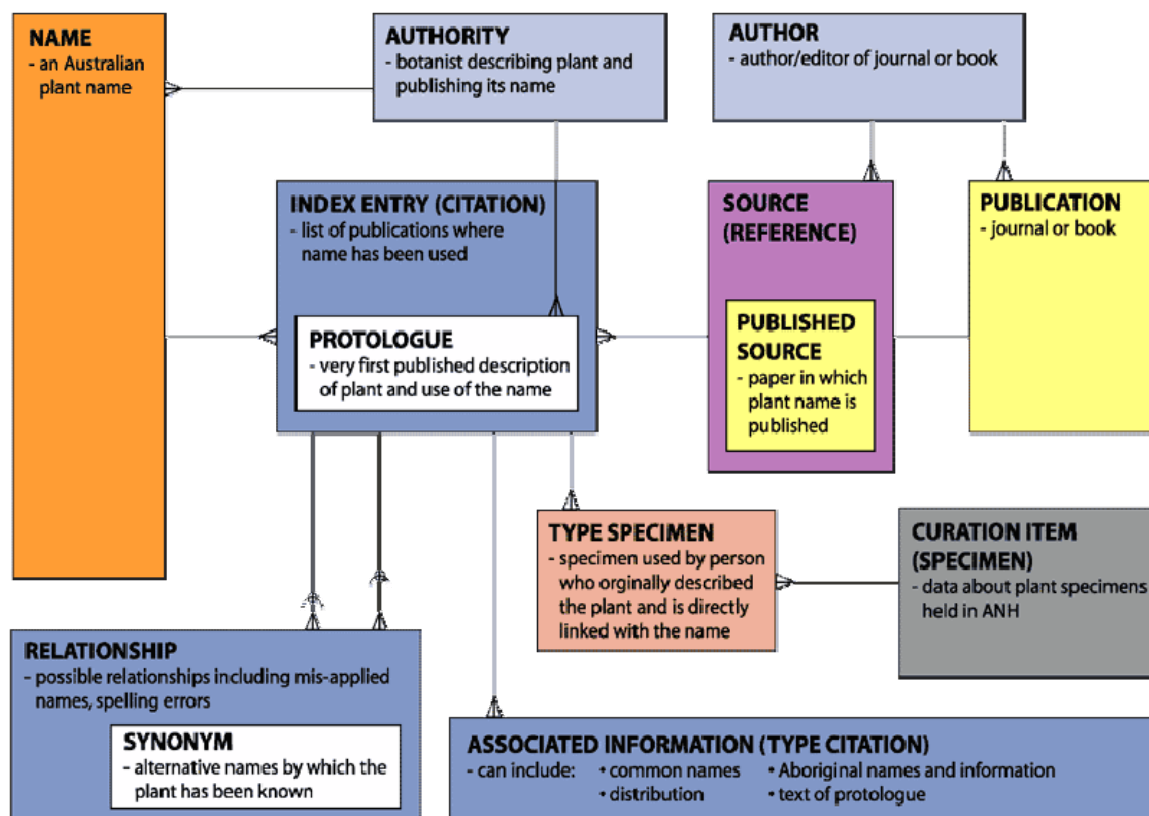
<sup>335</sup> <http://www.anbg.gov.au/databases/apni-about/index.html>

<sup>336</sup> <http://www.anbg.gov.au/databases/apni-about/index.html>

<sup>337</sup> <http://www.anbg.gov.au/cpbr/databases/apni.html>

<sup>338</sup> <http://www.anbg.gov.au/databases/apni-about/index.html>





While APNI offers a technical presentation of Australia's botanical nomenclature to the scientific community, *What's Its Name (WIN)* was developed as an alternative user-friendly entry point for the non-botanist. *What's Its Name* is derived from the APNI database and aims to deliver a concise overview of plant names and name changes via the Internet. It provides the current name in use for a plant and a cross-reference to any previously used names. In cases of discrepancy or ambiguity WIN provides a simple explanation of the scientific information suitable for the non-scientific user.<sup>339</sup>

APNI is involved with IPNI, an international collaboration between The Royal Botanic Gardens, Kew, The Harvard University Herbaria, and the Australian National Herbarium.<sup>340</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface
  - Personal and institutional use
  - Data is available tabulated or without

### Why use this tool?

- For an authoritative compilation of plant names<sup>341</sup>

### Who will use this tool?

- Data creation

<sup>339</sup> <http://www.anbg.gov.au/databases/apni-about/index.html>

<sup>340</sup> <http://en.wikipedia.org/wiki/IPNI>

<sup>341</sup> <http://www.anbg.gov.au/databases/apni-about/index.html>

- Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
- Data users
  - Expert
  - Interest groups
  - General public

#### **How will the tool be used?**

- Online searchable database
- The web versions of *APNI* are not case-sensitive and a trailing wild-card is assumed.<sup>342</sup>
- Pattern matching is based on the SQL database conventions<sup>343</sup>
- User input is required

#### **Where in the data chain could this tool be used?**

- Data source
- ALA central
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request
- As a post process, after data is with the user

#### **Availability**

- APNI: <http://www.anbg.gov.au/cpbr/databases/apni.html>
- ANBG: <http://www.anbg.gov.au/anbg/>
- WIN - What's Its Name: <http://www.anbg.gov.au/win/index.html>
- IPNI - International Plant Names Index: <http://www.ipni.org/>
- Actively maintained, available on the internet since 1991<sup>344</sup>
- Copyright various governments, institutions and individuals:  
<http://www.anbg.gov.au/copyright.html>
- Restrictions: personal, non-commercial use<sup>345</sup>
- Free access

#### **Comments**

- Scientific names only. No reference to common names.
- See [Australian Plant Census](#) for a national consensus of currently accepted names

---

<sup>342</sup> <http://www.anbg.gov.au/cpbr/databases/apni.html>

<sup>343</sup> <http://www.anbg.gov.au/cpbr/databases/apni.html>

<sup>344</sup> <http://www.anbg.gov.au/databases/apni-about/index.html>

<sup>345</sup> <http://www.anbg.gov.au/copyright.html>

## Australian Insect Common Names

Summary	
Type of tool	List of names
Function	Taxonomy, curatorial
Online / Desktop	Online database
Computer infrastructure	Web browser
Development status	Operational. Version 1.53 June 2005
Time of use	Data preparation, data processing
Licence	Free. Copyright CSIRO and AFFA

**Australian Insect Common Names** is a link between common and scientific names with images of some of the organisms.

### Description

This website provides ready access to the correct scientific name of every insect or related creature for which there is a common (or vernacular) name in use in Australia. The site also enables the user to discover the common name or names used in Australia for a species for which the user knows only the scientific name. Species are also listed in family groupings. An index of commonly used abbreviations of authors' names has also been included. This index is intended to assist in the interpretation of abbreviations which may be encountered in entomological literature. It is recommended, however, that in present-day usage authors' names be quoted in full to avoid ambiguity.

While scientific nomenclature is governed by strict rules, vernacular nomenclature is not. Inevitably there will be differences of opinion over what constitutes an appropriate common name or over whether a particular common name is or is not in wide use. In preparing the lists which follow we have endeavoured to include common names which are used in conversation and in the literature. We have also taken the opportunity to weed out a few contrived or clumsy names which have appeared in earlier editions of the Handbook but which seem not to be in use. Few Aboriginal names have been included but we believe that such names would enhance future versions of this website.

We have included the common names of Australian butterflies listed by M. Braby in *The Butterflies of Australia* (2001), although with some rationalisation.<sup>346</sup>

This website is based on the *CSIRO Handbook of Australian Insect Names* - 6th edition, 1993. The database of names behind that Handbook has been updated and augmented to reflect taxonomic changes, new names and newly recorded species.

Images and large-scale distribution maps accompany the names of arthropods and their relatives.<sup>347</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface

<sup>346</sup> <http://www.ento.csiro.au/aicn/backg.htm>

<sup>347</sup> <http://www.ento.csiro.au/aicn/intro.htm>

- Personal use

#### **Why use this tool?**

- A reliable source of common names and scientific names of insects

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data users
  - Expert
  - Interest groups
  - General public

#### **How will the tool be used?**

- Online tool
- Web browser
- User input required

#### **Where in the data chain could this tool be used?**

- Data source
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- As a post process, after data is with the user

#### **Availability**

- AICN: <http://www.ento.csiro.au/aicn/>
- Common names: [http://www.ento.csiro.au/aicn/name\\_c/a\\_1.htm](http://www.ento.csiro.au/aicn/name_c/a_1.htm)
- Scientific names: [http://www.ento.csiro.au/aicn/name\\_s/b\\_1.htm](http://www.ento.csiro.au/aicn/name_s/b_1.htm)
- Systematic names: <http://www.ento.csiro.au/aicn/system/system.htm>
- Authors: [http://www.ento.csiro.au/aicn/author/4a\\_main.htm](http://www.ento.csiro.au/aicn/author/4a_main.htm)
- Version 1.53 June 2005
- Copyright CSIRO and AFFA: <http://www.ento.csiro.au/aicn/acknow.htm>
- CSIRO legal notice: <http://www.csiro.au/org/LegalNoticeAndDisclaimer.html>

#### **Comments**

This web site is set up as an alphabetic index, and has no apparent search or remote call function.

## Australian Plant Census

Summary	
Type of tool	Database of accepted Australian plant names
Function	Taxonomy, name resolution
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Under development
Time of use	Before data is sent to ALA. At time of user request.
Licence	

The **Australian Plant Census** is a database of accepted plant names for Australia.

### Description

This is an enhancement to, and an integral part of, the Australian Plant Name Index (APNI).<sup>348</sup>

It aims to deliver a single, continually updatable, authoritative, fully synonymic census of Australia's vascular flora.<sup>349</sup>

In 2004 the Council of Heads of Australasian Herbaria (CHAH) agreed to produce a new cooperative census of Australian vascular plants. This Census will:<sup>350</sup>

- be a consensus view, endorsed by CHAH, of the taxa recognised as occurring in Australia
- provide an overview of alternative taxonomies
- provide an indication of geographic distribution for each taxon at State/Territory level
- list at least the more well-known vernacular names for each taxon
- be maintained and updated in perpetuity by CHAH

The APC will attempt to deliver a list containing all kinds of synonyms: formal (latinised) names published under the ICBN; formal (latinised) names existing in the literature but not formally published (manuscript names in current use); misapplied names in at least major works on the Australian flora; basionyms even if foreign; and informal phrase names. 'Foreign' synonyms will not normally be included, except on a case-by-case basis for some key groups (e.g. weeds).<sup>351</sup>

### Function

- Data cleaning and manipulation
  - Data cleaning – spelling, misnaming
  - Data validating – taxonomy
- Taxonomy
  - Identification tools
- User interface
  - Personal and institutional use
  - Raw data and visual presentation

### Why use this tool?

<sup>348</sup> <http://www.anbg.gov.au/chah/apc/introduction.html>

<sup>349</sup> <http://www.anbg.gov.au/chah/apc/background-orchard-2005.html>

<sup>350</sup> <http://www.anbg.gov.au/chah/apc/introduction.html>

<sup>351</sup> <http://www.anbg.gov.au/chah/apc/background-orchard-2005.html>

The Australian Plant Census is a major enhancement for the existing APNI, providing researchers, land managers, conservation agencies, regulatory agencies, industry and individuals with a single, unified, authoritative, annotated inventory of the Australian flora. When combined with the existing bibliographic data of APNI, and linked to other databases such as Australia's Virtual Herbarium, the Australian Plant Census is part of an unrivalled continental-scale biodiversity resource.<sup>352</sup>

#### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators
- Data users
  - Expert
  - Interest groups
- ALA infrastructure

#### How will the tool be used?

- Online tool
- Current/accepted names are in roman type.
- 'sensu' refers to the taxonomic and nomenclatural concept adopted for the APC.
- Synonyms are in *italic* type.
- Not all families have been treated for the APC, those not yet treated rely on plant names recognised in the 1990 *Census of Australian Vascular Plants* and/or more recent data from the *Australian Plant Name Index* (APNI).
- Occasional comments about the status of the name are based on decisions endorsed by CHAH.
- Full references are given at the end of the web page.
- Links from plant names take you to the more extensive entry of the *Australian Plant Name Index* (APNI)<sup>353</sup>

#### Where in the data chain could this tool be used?

- Data source
- ALA central
- User's machine

#### When could this tool be used?

- Before data is made available to ALA
- While data is stored with ALA
- At the time of a user request

#### Availability

- Australian Plant Census: <http://www.chah.gov.au/apc>
- CHAH - Council of Heads of Australasian Herbaria: <http://www.anbg.gov.au/chah/index.html>
- APNI - Australian Plant Name Index: <http://www.anbg.gov.au/apni/index.html>
- AVH - Australia's Virtual Herbarium: <http://www.anbg.gov.au/avh/>
- Background: <http://www.anbg.gov.au/chah/apc/background-orchard-2005.html>

<sup>352</sup> <http://www.anbg.gov.au/chah/apc/introduction.html>

<sup>353</sup> <http://www.anbg.gov.au/chah/apc/index.html>

- Cost: free

### **Comments**

## Australian Plant Common Names

Summary	
Type of tool	Database of names
Function	Taxonomy
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Operational
Time of use	At time of user request
Licence	

**Australian Plant Common Names** enables a search for common names and scientific names.

### Description

**Common name search using "kangaroo"**

<i>Anigozanthos manglesii</i>	Kangaroo Paw, Green (1)
<i>Anigozanthos manglesii</i>	Kangaroo Paw, Red (1)
<i>Anigozanthos pulcherrimus</i>	Kangaroo Paw, Golden (1)
<i>Anigozanthos viridis</i>	Kangaroo Paw, Green (1)
<i>Anigozanthos</i>	Kangaroo Paw (0)
<i>Cissus antarctica</i>	Kangaroo Vine (13)
<i>Cissus antarctica</i>	Vine, Kangaroo (13)
<i>Eucalyptus anceps</i>	Kangaroo Island Mallee (10)
<i>Eucalyptus cneorifolia</i>	Kangaroo Island Gum (10)
<i>Eucalyptus cneorifolia</i>	Kangaroo Island Narrow leaved Mallee (3)
<i>Eucalyptus remota</i>	Kangaroo Island Mallee Ash (3)
<i>Melaleuca halmaturorum</i>	Kangaroo Island Honey myrtle (10)
<i>Microsorium diversifolium</i>	Kangaroo Fern (7)
<i>Solanum aviculare</i>	Apple, Kangaroo (13)
<i>Solanum aviculare</i>	Kangaroo Apple (13)
Syn. <i>Acacia armata</i>	Kangaroo Thorn (1)
<i>Themeda australis</i>	Kangaroo Grass (1)

354

This search tool uses the Common Names database of the Australian National Botanic Gardens. Incomplete or partial names can be used as search terms.

References are also provided.

### Function

- Taxonomy
- User interface

<sup>354</sup> <http://www.anbg.gov.au/common.names/>



- Personal use
- Raw data and visual presentation

#### **Why use this tool?**

- To determine the common name and scientific name of a plant.

#### **Who will use this tool?**

- Data creation
- Data capture
  - Curators- identification
- Data users
  - Interest groups
  - General public
- ALA infrastructure
- No special skills required

#### **How will the tool be used?**

- Online tool
- Text entry and display
- Results are referenced but not guaranteed<sup>355</sup>
- User input required

#### **Where in the data chain could this tool be used?**

- ALA central
- User's machine

#### **When could this tool be used?**

- While data is stored with ALA
- At the time of a user request

#### **Availability**

- Australian Plant Common Names: <http://www.anbg.gov.au/common.names/>
- ANBG – Australian National Botanic Gardens: <http://www.anbg.gov.au/anbg/index.html>
- Cost: free

#### **Comments**

This tool could be modified, or the database tapped into, to provide an option for common names and scientific names for data requested from, or returned by, the ALA.

---

<sup>355</sup> <http://www.anbg.gov.au/common.names/>

## CAAB – Codes for Australian Aquatic Biota

Summary	
Type of tool	Coding system and searchable database
Function	Taxonomy, curatorial
Online / Desktop	Online database
Computer infrastructure	Web browser
Development status	Established. Version 2.0 2000
Time of use	Data preparation, data processing
Licence	Free

**CAAB - Codes for Australian Aquatic Biota** - is a continuously maintained and expanding 8-digit coding system for aquatic organisms in the Australian region maintained by CSIRO Division of Marine and Atmospheric Research, Australia (CMAR).<sup>356</sup>

### Description

Initially developed to cover fishes and selected other organisms of research or commercial interest, CAAB has more recently been expanded to provide more comprehensive coverage of a number of aquatic groups, as information is available. CAAB is used by a large and expanding number of governmental and industry groups for marine biology and fishery purposes and has replaced almost all other national and locally designed coding systems.

CAAB currently contains codes and taxonomic information for the following aquatic organisms in the Australian region:

- over 4,500 codes for fishes (CAAB category 37), representing virtually all known marine and freshwater species in Australian waters;
- codes for other marine vertebrates (CAAB categories 39 - reptiles, 40 - birds, 41 - mammals), representing all currently recognised marine species in Australia;
- over 14,000 codes for marine invertebrates (CAAB categories 10 to 36) in Australian waters, including sponges, stony corals, echinoderms, commercially important crustaceans and molluscs, tunicates, and other taxa;
- codes for Australian seagrasses and mangroves (CAAB category 63)
- codes for a representative selection of Australian seaweeds and microalgae (CAAB categories 52-55 and 70).<sup>357</sup>

The CAAB database can be searched on-line and detailed reports can be generated on any taxon. Users can use this interface to match a code to a taxon name, see current or alternative scientific or common names for any taxon or code, and access a variety of on-line links as available in the database.<sup>358</sup>

### Function

- Identification tools

### Why use this tool?

- The CAAB codes are used to identify aquatic organisms.

<sup>356</sup> <http://www.marine.csiro.au/caab/>

<sup>357</sup> <http://www.marine.csiro.au/caab/>

<sup>358</sup> <http://www.marine.csiro.au/caab/caabsearch.htm>

- The CAAB database can be used to search by scientific name, common name, or taxon code, or produce a list of taxa in the database for any category.<sup>359</sup>

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - Interest groups
  - General public

#### **How will the tool be used?**

- Online
- User input required

#### **Where in the data chain could this tool be used?**

- Data source
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- As a post process, after data is with the user

#### **Availability**

- CAAB: <http://www.marine.csiro.au/caab/>
- Cost: Free
- Copyright: CSIRO <http://www.csiro.au/org/LegalNoticeAndDisclaimer.html>

#### **Comments**

---

<sup>359</sup> <http://www.marine.csiro.au/caab/>

## Catalogue of Life

Summary	
Type of tool	Searchable database
Function	A name thesaurus and downloadable database for taxonomy
Online / Desktop	Online
Computer infrastructure	Web-base queries
Development status	Operational. 1 million species in 2007
Time of use	At data preparation and at time of user request
Licence	Copyright Species 2000. Use of content is encouraged

**Species 2000 & ITIS Catalogue of Life** is a searchable list of species provided by many taxonomic databases.

### Description

The Species 2000 & ITIS Catalogue of Life is planned to become a comprehensive catalogue of all known species of organisms on Earth by the year 2011.<sup>360</sup>

The present Catalogue is compiled with sectors provided by 47 taxonomic databases from around the world. Many of these contain taxonomic data and opinions from extensive networks of specialists, so that the complete work contains contributions from more than 3,000 specialists from throughout the taxonomic profession. Species 2000 and ITIS teams peer review databases, select appropriate sectors and integrate the sectors into a single coherent catalogue with a single hierarchical classification.<sup>361</sup>

The goal is to list every distinct species in each group of organisms. At present, some groups are globally complete, some are represented by global sectors that are nearing completion, and others are represented by partial sectors. The global sectors, whether complete or not, are provided by selected, peer reviewed global species databases in the Species 2000 federation or by equivalent global sectors of ITIS. The partial sectors are supplied principally by ITIS (N America), but also Species Fungorum and the Australian Faunal Directory, with the result that N American species are sometimes the only species represented for these incomplete groups.<sup>362</sup>

The Catalogue is published as two products:<sup>363</sup>

- **Species 2000 & ITIS Catalogue of Life: 2007 Annual Checklist**

The Annual Checklist is published each year as a fixed edition that can be cited and used as a common catalogue for comparative purposes by many organisations. A copy is on the CD-ROM, which is distributed free of charge, and an identical copy is on the website at <http://www.catalogueoflife.org/annual-checklist/2007/>. Archived earlier editions are also available on the website.

- **Species 2000 & ITIS Catalogue of Life: Dynamic Checklist**

The Dynamic Checklist is a virtual catalogue operated on the Internet and available both for users and as an electronic web-service at [www.catalogueoflife.org](http://www.catalogueoflife.org). The Dynamic Checklist harvests taxonomic sectors and associated strands of hierarchical classification dynamically from the source databases across the internet. The Dynamic Checklist is presently less extensive than the Annual Checklist because fewer taxonomic sectors have been connected so far.

<sup>360</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_about\\_col.php](http://www.catalogueoflife.org/annual-checklist/info_about_col.php)

<sup>361</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_about\\_col.php](http://www.catalogueoflife.org/annual-checklist/info_about_col.php)

<sup>362</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_2007\\_checklist.php](http://www.catalogueoflife.org/annual-checklist/info_2007_checklist.php)

<sup>363</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_about\\_col.php](http://www.catalogueoflife.org/annual-checklist/info_about_col.php)

It differs in concept from the Annual Checklist in that:

- i. the taxonomic records may be updated and the catalogue changed more frequently than in the Annual Checklist and
- ii. the Dynamic Checklist contains additional regional species checklists (such as the Regional Checklist - Europe, effectively a Pan-European Species Checklist) not included in the Annual Checklist.

Functionality of the Annual Checklist<sup>364</sup>

- Species (and infraspecific taxa) can be located either by searching by name or by tracking down through the hierarchical classification.
- Searching by name can be done using accepted scientific name, synonym or common name. Automatic synonymic and common name indexing takes the user directly to the species under its accepted name. The search can use part names, or be restricted to complete words.
- Tracking down the tree or classification uses accepted names for taxa.
- The structure of the Annual Checklist database has been optimised for performance with the user interface but is not ideal for importing to other systems.

Each species is listed with an accepted scientific name, a cited reference and its family and/or position in the hierarchical classification. Additional common names and synonyms may be provided, but these data are not complete, and for some species none may exist.<sup>365</sup>

Species 2000 and ITIS organisations work together to create the Catalogue of Life.

### ITIS

The Integrated Taxonomic Information System (ITIS) is a partnership of federal agencies and other organisations from the United States, Canada, and Mexico, with data stewards and experts from around the world (see <http://www.itis.gov>). The ITIS database is an automated reference of scientific and common names of biota of interest to North America. It contains more than 335,000 names for species in all kingdoms.<sup>366</sup>

### Species 2000

Species 2000 (<http://www.sp2000.org>) is an autonomous federation of taxonomic database custodians, involving taxonomists throughout the world. Its goal is to collate a uniform and validated index to the world's known species. There are two regional programmes: Species 2000 europa (<http://www.sp2000europa.org>) and Species 2000 Asia-Oceania ([www.sp2000asiaoceania.org](http://www.sp2000asiaoceania.org)).<sup>367</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface
  - Personal and institutional use
  - Raw data and web-based visual presentation

### Why use this tool?

For authoritative taxonomic information on common names and scientific names

---

<sup>364</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_2007\\_checklist.php](http://www.catalogueoflife.org/annual-checklist/info_2007_checklist.php)

<sup>365</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_2007\\_checklist.php](http://www.catalogueoflife.org/annual-checklist/info_2007_checklist.php)

<sup>366</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_about\\_col.php](http://www.catalogueoflife.org/annual-checklist/info_about_col.php)

<sup>367</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_about\\_col.php](http://www.catalogueoflife.org/annual-checklist/info_about_col.php)

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - Interest groups
  - General public
- ALA infrastructure

### How will the tool be used?

- Web based online tool
- Results can be exported as an Excel spreadsheet
- User input is required

### Where in the data chain could this tool be used?

- Data source
- ALA central – if the Annual Checklist database is used
- User's machine
- Pathways between these

### When could this tool be used?

- Before data is made available to ALA
- While data is stored with ALA
- At the time of a user request

### Availability

- Catalogue of Life: <http://www.catalogueoflife.org/>
- Search: <http://www.catalogueoflife.org/annual-checklist/search.php>
- Browse taxonomic tree: [http://www.catalogueoflife.org/annual-checklist/browse\\_taxa.php](http://www.catalogueoflife.org/annual-checklist/browse_taxa.php)
- Contact: [http://www.catalogueoflife.org/annual-checklist/info\\_contact\\_us.php](http://www.catalogueoflife.org/annual-checklist/info_contact_us.php)
- Copyright Species 2000

Use of the content (such as the classification, synonymic species checklist, and scientific names) for publications and databases by individuals and organisations for not-for-profit usage is encouraged, on condition that full and precise credit is given at three levels on all occasions that records are shown. The three levels credit the complete work, the contributing database of the record, and the expert who provides taxonomic scrutiny of the individual record.<sup>368</sup>

### Comments

An example of a search of Catalogue of Life: Gummy Shark - [http://www.catalogueoflife.org/annual-checklist/search\\_results.php?search\\_string=gummy+shark](http://www.catalogueoflife.org/annual-checklist/search_results.php?search_string=gummy+shark)  
This is a largely European and American catalogue and its coverage of Australian species is limited.

---

<sup>368</sup> [http://www.catalogueoflife.org/annual-checklist/info\\_copyright.php](http://www.catalogueoflife.org/annual-checklist/info_copyright.php)

## CAVS Biocodes – Census of Australian Vertebrate Species

Summary	
Type of tool	Coding system and searchable database
Function	Taxonomy, curatorial
Online / Desktop	Online database
Computer infrastructure	Web browser
Development status	
Time of use	Data preparation, data processing
Licence	Free

### Description

In the Australian fauna, some groups have been given Biocodes, such as the *Census of Australian Vertebrate Species* (CAVS) for Amphibia, Aves, Mammalia (Eutheria, Marsupialia, Prototheria) and Reptilia and Codes for Australian Aquatic Biota (CAAB) for aquatic organisms, mainly fishes.<sup>369</sup>

The CAVS searchable database contains codes and information on distribution, taxonomy, ecology and bibliography, and links to the Australian Faunal Directory.

Part of the ABRS Australian Biological Resources Study.

### Function

- Identification tools
  - keys

### Why use this tool?

- Taxonomy and identification

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators - identification
- Data users
  - Expert
  - Interest groups

### How will the tool be used?

- Online tool
- User input required – biocode or name

### Where in the data chain could this tool be used?

- Data source
- User's machine

### When could this tool be used?

- Before data is made available to ALA

---

<sup>369</sup> <http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/cavs/>

- At the time of a user request
- As a post process, after data is with the user

**Availability**

- CAVS: <http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/cavs/>
- ABRS Fauna Online: <http://www.environment.gov.au/biodiversity/abrs/online-resources/fauna/>
- Copyright © Commonwealth of Australia, Department of Environment and Water Resources: <http://www.environment.gov.au/about/copyright.html>

**Comments**



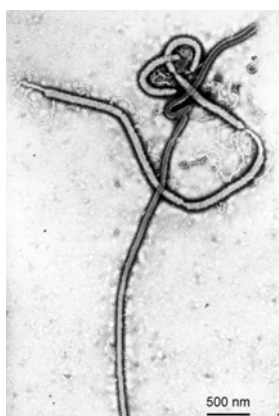
## ICTVdB – Virus Database

Summary	
Type of tool	Database of viruses
Function	Virus taxonomy
Online / Desktop	Online with desktop identification key
Computer infrastructure	Browser. Identification using Intkey software.
Development status	Actively maintained and developed
Time of use	Data creation. As a post process on users machine.
Licence	

The **Universal Virus Database of the International Committee on Taxonomy of Viruses**. The **ICTVdB** is a dynamic database containing information about viruses of animals, plants, bacteria, and fungi.<sup>370</sup>

### Description

The directory of ICTVdB is an Index of Viruses, a list of approved virus names linked to virus descriptions coded from information in Virus Taxonomy. It also incorporates the plant virus database VIDeB and is illustrated with Electron Microscope pictures, diagrams and images of symptoms contributed by virologists around the world.<sup>371</sup>



*Ebola virions, diagnostic specimen from the first passage in Vero cells of a specimen from a human patient — this image is from the first isolation and visualization of Ebola virus, 1976. In this case, some of the filamentous virions are fused together, end-to-end, giving the appearance of a "bowl of spaghetti."*<sup>372</sup>

ICTVdB uses a decimal code to present, on the Web:

- searchable descriptions of virus isolates, species, genera, families, orders,
- images of many viruses, and
- links to genomic and protein databanks.

ICTVdB uses DELTA format:

- for user-friendly simultaneous access to character list, data files and specification files
- for interactive identification and data retrieval,
- for data conversion shopfronts eg for phylogenetic research.

<sup>370</sup> [http://www.ncbi.nlm.nih.gov/ICTVdb/Tutorial/tt\\_intro.htm](http://www.ncbi.nlm.nih.gov/ICTVdb/Tutorial/tt_intro.htm)

<sup>371</sup> <http://www.ncbi.nlm.nih.gov/ICTVdb/canintro1.htm>

<sup>372</sup> <http://www.ncbi.nlm.nih.gov/ICTVdb/Images/Murphy/ebola2.htm>

ICTVdB is also using:

- entry forms on the Web to peer review new virus data through ICTV Study Groups.
- descriptions on the level of isolates and species for all known viruses.<sup>373</sup>

The character list for ICTVdB consists of over 2600 different viral properties which are used to describe and characterise viral taxa.<sup>374</sup> All virus descriptions are based on the character list and natural language translations from the encoded descriptions are automatically generated and formatted for display.<sup>375</sup>

Initially designed for taxonomic research, the ICTVdB has evolved to become a major reference resource and research tool.<sup>376</sup>

### Function

- Visualisation tools
  - Images
- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Raw data and visual presentation

### Why use this tool?

- An authoritative source for virus names and descriptions

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data providers
  - Institutions
- Data users
  - Expert

### How will the tool be used?

To use the identification and information retrieval program, the Intkey software and Intkey data files are needed. The Intkey software is freely available.<sup>377</sup>

- Online searchable database
- Virus descriptions are generated automatically from DELTA files
- User input is required.

### Where in the data chain could this tool be used?

- Data source
- User's machine

### When could this tool be used?

- Before data is made available to ALA

---

<sup>373</sup> <http://www.ncbi.nlm.nih.gov/ICTVdb/canintro1.htm>

<sup>374</sup> <http://phene.cpmc.columbia.edu/charsindex.htm>

<sup>375</sup> <http://phene.cpmc.columbia.edu/ICTVdB/48102001.htm>

<sup>376</sup> [http://www.ncbi.nlm.nih.gov/ICTVdb/Tutorial/tt\\_intro.htm](http://www.ncbi.nlm.nih.gov/ICTVdb/Tutorial/tt_intro.htm)

<sup>377</sup> <http://www.ncbi.nlm.nih.gov/ICTVdb/WIntkey/index0.htm>

- At the time of a user request
- As a post process, after data is with the user

#### **Availability**

- ICTVdB: <http://www.ncbi.nlm.nih.gov/ICTVdb/index.htm>
- Index of viruses: <http://www.ncbi.nlm.nih.gov/ICTVdb/Ictv/ICTVindex.htm>
- Tutorial: <http://www.ncbi.nlm.nih.gov/ICTVdb/Tutorial/>
- Intkey required for identification: <http://delta-intkey.com/>
- Actively developed and maintained by [Dr Cornelia Büchen-Osmond](#)
- Copyright: International Committee on Taxonomy of Viruses
- Licence not mentioned

#### **Comments**

- Some descriptions are only very basic and links may point to documents that are not yet published on the Web.<sup>378</sup>

---

<sup>378</sup> <http://phene.cpmc.columbia.edu/ICTVdB/48102001.htm>

## Index Fungorum

Summary	
Type of tool	Database of fungi
Function	Fungus taxonomy
Online / Desktop	Online database
Computer infrastructure	Web browser
Development status	Active and current
Time of use	At time of user request
Licence	Community resource. IP lies with contributors

**Index Fungorum** is the world database of fungal names (including yeasts, lichens, chromistan fungi, protozoan fungi and fossil forms) at species level and below.

### Description

The Index of Fungi is a publication from CABI and currently provides the majority of recently published names. Certain data elements (name string, author string, year of publication) from the Index of Fungi are made available through Index Fungorum. The remaining data elements are not immediately available but are made so after five years.<sup>379</sup>



*Amanita muscaria*<sup>380</sup>

For most entries the taxonomic hierarchy attached to a name is either derived from the nomenclatural position of the genus through its type (thus *Helotium* appears in the Tricholomataceae because the type of *Helotium* is a basidiomycete), or in some cases through the presumed position based on an assessment of associated homotypic names. Only in GSD (Global Species Database) data, or other data for which a taxonomic opinion is available (and which contribute to the Catalogue of Life - see Species 2000), is the correct taxonomic position of a name known through the position of the currently accepted name.

<sup>379</sup> <http://www.indexfungorum.org/Names/Names.asp>

<sup>380</sup> <http://www.rbgsvd.nsw.gov.au/>

The database has been derived from a number of published lists including Saccardo's Sylloge Fungorum, Petrak's Lists, Saccardo's Omissions, Lamb's Index, Zahlbruckner's Catalogue of Lichens (comprehensive for names at species level only but with an increasing number of names of infraspecific taxa) and the Index of Fungi. A name record will usually have a reference to an entry in one or more of these published lists in addition to other data derived from numerous acknowledged sources. Author citations conform to the Brummitt & Powell standard.<sup>381</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

- An authoritative source of fungi names.

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data users
  - Expert
  - Interest groups

### How will the tool be used?

Persistent identifiers have been implemented for the records in the Index Fungorum database. Each record is assigned an LSID with the following structure:<sup>382</sup>

**urn:lsid:indexfungorum.org:names:nnnnnn**

The current database structure does not support the assignment of LSIDs to the taxonomic opinions expressed in the 'Species Fungorum' component; only the names in a nomenclatural sense are assigned LSIDs.

Search terms can be truncated without using wildcards. The database search restricts you to 6000 records so if you enter something as vague as 'A' you will **not** get all the records.

Names with **blue** links are names for which there is no taxonomic opinion available, **red** links are misapplied names (names which have been used in a sense which is different from that as represented by the type of the name), **green** links are names where a taxonomic opinion has been expressed and lead directly to the appropriate page.

Entries with '[GSD]' or '[RSD]' have onward links to Global Species Databases or Regional Species Databases and full synonymy data.<sup>383</sup>

Full publication details are provided for some entries from CABI's Index of Fungi. However, full details of names published during the last 5 years have been omitted<sup>384</sup> to encourage you to take out a subscription.

---

<sup>381</sup> <http://www.indexfungorum.org/Names/Names.asp>

<sup>382</sup> <http://www.indexfungorum.org/Names/Names.asp>

<sup>383</sup> <http://www.indexfungorum.org/Names/Names.asp>

<sup>384</sup> <http://www.indexfungorum.org/Names/Names.asp>

- Online searchable database
- Cookies should be enabled for navigation
- User input required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- At the time of a user request

**Availability**

- CABI Bioscience Databases: <http://www.indexfungorum.org/>
- Search: <http://www.indexfungorum.org/Names/Names.asp>
- Development is active and current. 400,000 records online.
- Cost: free
- Intellectual property lies with the contributors<sup>385</sup>

**Comments**

- No common names
- Commonly there may be an image of the page of the original citation

---

<sup>385</sup> <http://www.indexfungorum.org/Names/Names.asp>

## IPNI – International Plant Names Index

Summary	
Type of tool	Searchable database
Function	Taxonomy, curatorial
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Active
Time of use	Data preparation, data processing
Licence	

**IPNI - International Plant Names Index** is an international searchable database of plant names.

### Description

The International Plant Names Index (IPNI) is a database of the names and associated basic bibliographical details of all seed plants, ferns and fern allies. Its goal is to eliminate the need for repeated reference to primary sources for basic bibliographic information about plant names. The data are freely available and are gradually being standardized and checked. IPNI will be a dynamic resource, depending on direct contributions by all members of the botanical community.<sup>386</sup>

IPNI also acts as an LSID server, allowing the resolution of specially formatted IPNI LSIDs automatically into RDF format metadata.<sup>387</sup>

IPNI is the product of a collaboration between The Royal Botanic Gardens, Kew, The Harvard University Herbaria, and the Australian National Herbarium.<sup>388</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface
  - Personal and institutional use
  - Raw data or visual presentation

### Why use this tool?

- For an authoritative compilation of plant names

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
- Data users

---

<sup>386</sup> <http://www.ipni.org/index.html>

<sup>387</sup> <http://www.ipni.org/lisids.html>

<sup>388</sup> <http://www.ipni.org/index.html>

- Expert
- Interest groups
- General public

#### **How will the tool be used?**

IPNI data can be downloaded in a number of different delimited data formats, as well as the normal and full record html formats. The normal format gives you the classic IPNI listing, 100 records per page. The full record format gives you a listing but with the full details displayed for each record.<sup>389</sup>

- Online database
- Web access
- User input required

#### **Where in the data chain could this tool be used?**

- Data source
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request
- As a post process, after data is with the user

#### **Availability**

- IPNI: <http://www.ipni.org/>
- Linking to IPNI: [http://www.ipni.org/link\\_to\\_ipni.html](http://www.ipni.org/link_to_ipni.html)
- Data formats: [http://www.ipni.org/ipni/delimited\\_help.html](http://www.ipni.org/ipni/delimited_help.html)
- Free for personal or use within an organisation
- Copyright Plant Names Project<sup>390</sup>
- Restrictions: Max download 5000 records

#### **Comments**

- Search can be slow
- Scientific names. No common names
- Lacks information on many species

---

<sup>389</sup> [http://www.ipni.org/ipni/delimited\\_help.html](http://www.ipni.org/ipni/delimited_help.html)

<sup>390</sup> <http://www.ipni.org/copyright.html>



## ITIS - Integrated Taxonomic Information System

Summary	
Type of tool	Searchable database
Function	A name thesaurus and downloadable database for taxonomy
Online / Desktop	Online
Computer infrastructure	Web-base queries
Development status	Operational. 460,000 scientific names November 2007
Time of use	At data preparation and at time of user request
Licence	Public information for distribution and copying

ITIS is authoritative taxonomic information on plants, animals, fungi, and microbes.<sup>391</sup>

### Description

ITIS is a searchable and downloadable database of common and scientific names and their taxonomy.

An overriding goal of the ITIS project is to provide accurate, scientifically credible, and current taxonomic data that meet the needs of the ITIS partners and the user public.<sup>392</sup>

**Quick search** allows searching on the taxon name, either common or scientific.

**Advanced Search and Report** allows for:

- search on taxon author
- filter on Record Credibility Rating
- select fields to display in report output<sup>393</sup>

ITIS does not intend to serve as a forum for cutting-edge taxonomic classifications. Rather, ITIS is meant to serve as a standard to enable the comparison of biodiversity datasets, and therefore aims to incorporate classifications that have gained broad acceptance in the taxonomic literature and by professionals who work with the taxa concerned.

### Data Standards

- Data conform to the **International Code of Botanical Nomenclature** and the **International Code of Zoological Nomenclature**.
- A **five kingdom system** has been adopted as a standard. The kingdoms are *Monera*, *Protista*, *Plantae*, *Fungi*, and *Animalia*.
- **Ranks** in the animal kingdom below *subspecies* will not be included in ITIS regardless of their occasional inclusion in datasets, as these ranks are not allowed under the zoological code. The botanical code allows the ranks *variety*, *subvariety*, *forma*, and *subforma*.
- **Protists**, depending on the group considered, have historically been treated according to either the zoological code (i.e. protozoans) or the botanical code (i.e. algae) and in some cases according to both. Also, the algae are often classified into multiple kingdoms. ITIS will make **practical** decisions as to the placement of these groups within the five kingdom framework.<sup>394</sup>

---

<sup>391</sup> <http://www.itis.gov/>

<sup>392</sup> <http://www.itis.gov/standard.html>

<sup>393</sup> [http://www.itis.gov/advanced\\_search.html](http://www.itis.gov/advanced_search.html)

<sup>394</sup> <http://www.itis.gov/standard.html>

We are a partnership of US, Canadian, and Mexican agencies ([ITIS-North America](#)); other organizations; and taxonomic specialists. ITIS is also a partner of [Species 2000](#) and the [Global Biodiversity Information Facility \(GBIF\)](#).<sup>395</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface
  - Personal and institutional use
  - Raw data and web-based visual presentation

### Why use this tool?

For authoritative taxonomic information on common names and scientific names

### Who will use this tool?

This tool could be used by a wide variety of users:

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - Interest groups
  - General public
- ALA infrastructure

### How will the tool be used?

ITIS has two forms of data access – Search and Report, and Download Database.<sup>396</sup>

#### Search and Report

- **Advanced Search and Report** - results in a customized report including only the taxon categories chosen from the ITIS database.
- **Hierarchical Report** - view a listing, from a single kingdom, of scientific names in hierarchical order. This report includes point-and-click retrieval of information associated with each scientific name in the output.

#### Download Database

- **ITIS Customized Download** - Create a download file for any part of the ITIS database based upon specifications you define.
  - **Download a specific taxonomic group** - Create a download file for a specific taxonomic group.
  - The full ITIS database may also be downloaded.
- This information is accessed online through a web browser

---

<sup>395</sup> <http://www.itis.gov/index.html>

<sup>396</sup> <http://www.itis.gov/access.html>

- The entire database is available via FTP
- User input is required

#### **Where in the data chain could this tool be used?**

- Data source
- ALA central - if, for example, the full ITIS database were used
- User's machine
- Pathways between these

#### **When could this tool be used?**

- Before data is made available to ALA
- As data is imported into ALA for storage
- At the time of a user request

#### **Availability**

- ITIS search: <http://www.itis.gov/>
- Instructions for linking directly to the ITIS Standard Report: [http://www.itis.gov/standard\\_report\\_link.html](http://www.itis.gov/standard_report_link.html)
- Guidelines for data submission: <http://www.itis.gov/submit.html>
- Information presented on the ITIS website is considered public information and may be distributed or copied.

#### **Comments**

ITIS have several tools for taxonomists:<sup>397</sup>

- **Compare Taxonomy/Nomenclature.** An application to match your list of taxa to the ITIS Database.
- **The Taxonomic Workbench** is a Windows-based software tool used for editing and manipulating taxonomic data for submission into the ITIS online system.

---

<sup>397</sup> <http://www.itis.gov/tools.html>

## LPSN – List of Prokaryotic names with Standing in Nomenclature

Summary	
Type of tool	Database of bacteria
Function	Bacteria taxonomy
Online / Desktop	Online database
Computer infrastructure	Browser
Development status	Active. Recently and regularly updated
Time of use	At time of user request
Licence	Personal and non-commercial use

**List of Prokaryotic names with Standing in Nomenclature** includes, alphabetically and chronologically, the nomenclature of bacteria and the nomenclatural changes as cited in the *Approved Lists of Bacterial Names* or validly published in the *International Journal of Systematic Bacteriology* (IJSB) or in the *International Journal of Systematic and Evolutionary Microbiology* (IJSEM).<sup>398</sup>

### Description

There is no official classification of prokaryotes, but the names given to prokaryotes are regulated. The *International Code of Nomenclature of Bacteria (Bacteriological Code)* contains General Considerations, Principles, Rules and Recommendations which govern the way in which the names of prokaryotes are to be used.<sup>399</sup>

The citations of names are in the correct format according to the *Bacteriological Code* (1990 Revision), and the nomenclatural types are provided. Abbreviations and addresses of collections from which designated type strains are available are listed alphabetically. Basonyms and synonyms are included to clarify the previous names or histories of individual taxa.<sup>400</sup>

The names in this list are "valid" only in the sense of being validly published as a result of conformity with the Rules of Nomenclature. The inclusion of a name on this list is not to be construed as taxonomic acceptance of the taxon to which the name is applied. Indeed, some of these names may, in time, be shown to be synonyms, or the organisms may be transferred to another genus, thus necessitating the creation of a new combination.<sup>401</sup>

### Function

- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Raw data or visual presentation

### Why use this tool?

- A search for the names of bacteria and other prokaryotic names

### Who will use this tool?

- Data creation

---

<sup>398</sup> <http://www.bacterio.cict.fr/generalinformation.html>

<sup>399</sup> <http://www.bacterio.cict.fr/introduction.html>

<sup>400</sup> <http://www.bacterio.cict.fr/introduction.html>

<sup>401</sup> <http://www.bacterio.cict.fr/introduction.html>

- Experts - taxonomy
- Data users
  - Expert
  - Interest groups

**How will the tool be used?**

- Online searchable database
- User input is required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- At the time of a user request

**Availability**

- LPSN: <http://www.bacterio.net> or <http://www.bacterio.cict.fr/>
- Search: <http://www.bacterio.cict.fr/search.html>
- All names in LSPN: <http://www.bacterio.cict.fr/allnames.html>
- Active since March 1997. Updated with the publication of each new issue of the *International Journal of Systematic and Evolutionary Microbiology*.<sup>402</sup>
- Free access
- Copyright: J.P. Euzéby
- Licence: Personal and non-commercial use.
- Restrictions: No copying, distribution, altering <http://www.bacterio.cict.fr/copyright.html>

**Comments**

One of several databases of bacteria names.

---

<sup>402</sup> <http://www.bacterio.cict.fr/introduction.html>

## TROPICOS

Summary	
Type of tool	Database of plant names
Function	Taxonomy
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Revision 1.5
Time of use	Data preparation, at time of user request, post process
Licence	Non-commercial, not for publication, no uncontrolled subsets

**TROPICOS**, the world's largest database of plant information, contains fully web-searchable records for over 900,000 plant names and nearly 2 million specimens. Over 50,000 plant images are also linked to their records in TROPICOS.<sup>403</sup>

### Description

W<sup>3</sup>TROPICOS provides new and improved access to the Missouri Botanical Garden's VAST (VAScular Tropicos) nomenclatural database and associated authority files. The following information is provided when present:<sup>404</sup>

#### Names data

- Plant name and authors
- Group and family placement
- Place and date of publication
- Type information
- Basionym, with place and date of publication
- Next Higher Taxon, with place and date of publication
- Other uses of this name
- Synonyms of this name, and References for the alternate usage
- Homonyms and Intraspecific names for species

#### Reference

- Author(s) of the publication
- Date of publication
- Title of the article
- Journal or book title
- Volume and page numbers
- Keywords

### Function

- Data cleaning and manipulation
  - Data cleaning – spelling, misnaming
  - Data validating – taxonomy, geography
- Taxonomy
  - Identification tools, keys
- User interface
  - Personal use
  - Visual presentation

---

<sup>403</sup> <http://www.mobot.org/plantscience/>

<sup>404</sup> <http://mobot.mobot.org/W3T/Search/vast.html>

**Why use this tool?**

- Correct plant name and synonym
- Plant images

**Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- Data users
  - Expert
  - Interest groups
- Scientific names only

**How will the tool be used?**

- Online database of names
- User input is required
- Not a batch job

**Where in the data chain could this tool be used?**

- Data source
- User's machine

**When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request
- As a post process, after data is with the user

**Availability**

- TROPICOS: <http://mobot.mobot.org/W3T/Search/vast.html>
- Web TROPICOS (beta): <http://test.tropicos.org/>
- Missouri Botanical Garden: <http://www.mobot.org/>
- Cost: free
- Revision 1.5
- Copyright Missouri Botanical Garden
- Restrictions: Non-commercial, not for republication. No uncontrolled subsets.

**Comments**

- May not be the most complete source for Australian plants.
- No common names

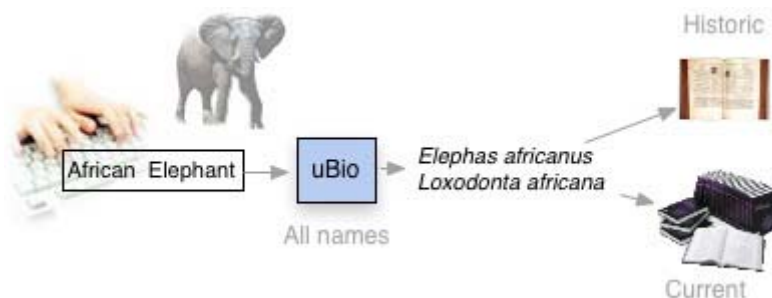
## uBio Taxonomic Name Server

Summary	
Type of tool	Set of tools
Function	A name thesaurus for taxonomy
Online / Desktop	Online
Computer infrastructure	Protocols used are SOAP and LSIDs
Development status	Developmental and operational, 10 million names in November 2007
Time of use	At the time of a user request
Licence	Academic Free Licence

**uBio - Universal Biological Indexer and Organizer** uses names and taxonomic intelligence to manage information about organisms.<sup>405</sup>

### Description

uBio is an initiative within the science library community to join international efforts to create and utilize a comprehensive and collaborative catalog of known names of all living (and once-living) organisms. The Taxonomic Name Server (TNS) catalogs names and classifications to enable tools that can help users find information on living things using any of the names that may be related to an organism.<sup>406</sup>



Information about organisms is often linked to a name.

This can create problems in information retrieval because:

- One taxon can have many names.
- The same name can refer to many taxa.

uBio is working on tools for providers of biological information that address these problems.

The uBio **Taxonomic Name Server** acts as a name thesaurus.

Names have many different classes of relationships that can be used to organize and retrieve information that is annotated with names. These classes are divided into two inter-connected services.

**NameBank** is a repository of millions of recorded biological names and facts that link those names together.

**ClassificationBank** stores multiple classifications and taxonomic concepts that are the result of expert opinions. It extends the functionality of NameBank.

<sup>405</sup> <http://www.ubio.org/>

<sup>406</sup> <http://www.ubio.org/index.php?pagename=general>



All data within these components are linked to mechanisms that provide credit and attribution to experts who provide name and linkage information within the TNS.

Lastly, NameBank promotes the emergence of a layered biological informatics infrastructure that allows different expert systems to share common information. This conserves scarce resources and enhances the means to support continued expert work.<sup>407</sup>

Some of the tools algorithms and services developed by uBio:<sup>408</sup>

- **LinkIT - Dynamic Linking** - automatically link names in your site to authoritative databases.
- **FindIT - Name Recognition Tool** - identifies scientific names in any text.
- **Author Abbreviation Resolver** is a thesaurus for resolving abbreviations of author names in scientific nomenclature.
- **ParseIT - Deconstruct a scientific name** This tool accepts a complex scientific name and breaks it into its component parts.
- **CrawlIT -Find all the names in a web site.** Locate all the names within a collection of content and match the results against NameBank and various authority lists.
- **X:ID - Build your own identification keys.** This identification key software is for the creation and display of interactive taxonomic keys. It is written entirely in OpenSource code and is XML-based.
- **CompareIT** takes a URL or a list of names as input and compares the taxon names with a current taxonomy such as Species 2000 or ITIS and reports on the current status of the name and other metrics.
- **uBioRSS** A taxonomically intelligent RSS aggregator. Two applications that draw upon text-matching algorithms in conjunction with RSS feeds provided by academic publishers.<sup>409</sup>

uBio has its origins in the MBLWHOI Library in Woods Hole, Massachusetts, part of the Marine Biological Laboratory. FishBase, AntWeb and botanicus are uBio-based Services.<sup>410</sup>

### Function

- Taxonomy

### Why use this tool?

Information in biology related to organisms is often annotated with a name. We refer to this content as name-bearing or *biocentric*, content.

Names would appear to be a logical candidate for keyword searches within data repositories. The only problem is that organism names are neither fixed, stable, or unique and employing them as query terms can result in receiving information that is not related to the organism you were looking for as well as missing information you wanted.<sup>411</sup>

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification

---

<sup>407</sup> [http://www.ubio.org/index.php?pagename=background\\_intro](http://www.ubio.org/index.php?pagename=background_intro)

<sup>408</sup> [http://www.ubio.org/index.php?pagename=sample\\_tools](http://www.ubio.org/index.php?pagename=sample_tools)

<sup>409</sup> <http://www.ubio.org/index.php?pagename=ubioRSS>

<sup>410</sup> <http://www.ubio.org/>

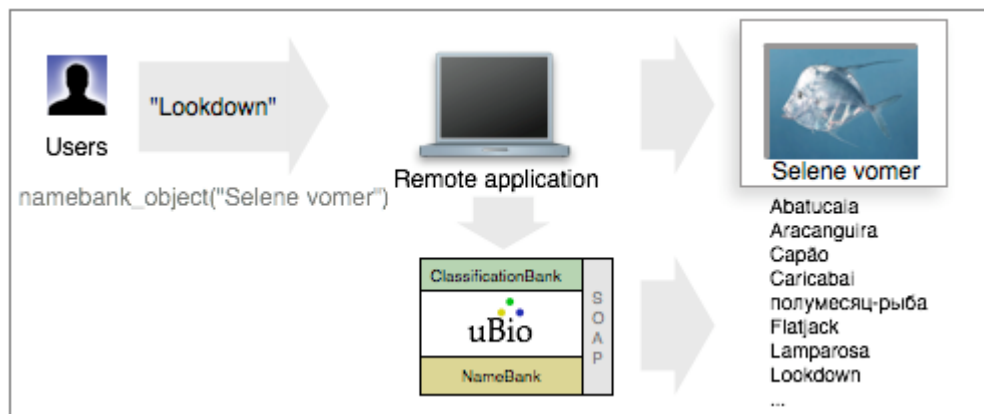
<sup>411</sup> [http://www.ubio.org/index.php?pagename=data\\_objects](http://www.ubio.org/index.php?pagename=data_objects)

- Data providers
  - Institutions
  - Private collections
- ALA infrastructure
- Data users
  - Expert
  - Interest groups

### How will the tool be used?

NameBank data are accessed via methods defined in SOAP. The SOAP methods provide access to lists of languages, management taxonomy "packages", and name search interfaces. The primary output object is the NameBank object. The NameBank object contains version and source metadata, record confidence and namestring metadata, and arrays of associated nominal, nomenclatural and canonical group identifiers and name strings. The next version of NameBank will move output formats to RDF/XML which will be resolved through LSIDs.<sup>412</sup>

The web services allows users to access uBio data as if it were a local resource. For example, a library may have a database of fish pictures it serves. Users may query by name to find pictures. The developer of this system could use NameBank to access additional names that can be used to ensure that name queries find the pictures even if the name wasn't originally attached to the picture.



A remote application uses uBio web services to add taxonomic intelligence to application. In this figure an application serving images annotated with scientific names uses uBio web services to allow users to find images when searching by common name.<sup>413</sup>

### Where in the data chain could this tool be used?

- Data source
- ALA central
- Pathways between these

### When could this tool be used?

- Before data is made available to ALA
- As data is imported into ALA for storage
- While data is stored with ALA

<sup>412</sup> [http://www.ubio.org/index.php?pagename=namebank\\_out](http://www.ubio.org/index.php?pagename=namebank_out)

<sup>413</sup> [http://www.ubio.org/index.php?pagename=services\\_overview](http://www.ubio.org/index.php?pagename=services_overview)

- At the time of a user request

#### **Availability**

- uBio: <http://www.ubio.org/>
- Sample applications: [http://www.ubio.org/index.php?pagename=sample\\_applications](http://www.ubio.org/index.php?pagename=sample_applications)
- Tools and services: [http://www.ubio.org/index.php?pagename=sample\\_tools](http://www.ubio.org/index.php?pagename=sample_tools)
- uBio news RSS: <http://www.ubio.org/apps/uBioNews/rss.php>
- Software is copyright MBL/WHOI Library and licensed under the Academic Free License<sup>414</sup>

#### **Comments**

- Used by the Biodiversity Heritage Library
- uBio has a strong North American representation, and may have fewer Australian species.

---

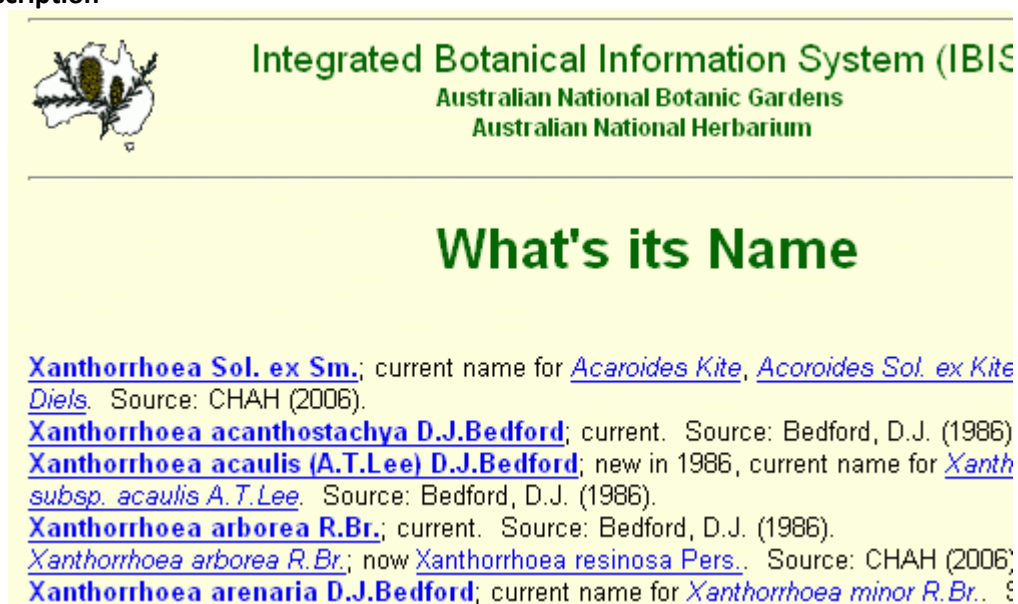
<sup>414</sup> [http://www.ubio.org/index.php?pagename=soap\\_methods/namebank\\_search](http://www.ubio.org/index.php?pagename=soap_methods/namebank_search)

## WIN? – What's Its Name?

Summary	
Type of tool	Database of plant names
Function	Taxonomy
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Operational
Time of use	At time of user request
Licence	

**What's Its Name?** is a concise database of plant names and name changes for Australia for the non-botanist.

### Description



**Integrated Botanical Information System (IBIS)**  
Australian National Botanic Gardens  
Australian National Herbarium

## What's its Name

[Xanthorrhoea Sol. ex Sm.](#); current name for [Acaroides Kite](#), [Acoroides Sol. ex Kite Diels.](#) Source: CHAH (2006).  
[Xanthorrhoea acanthostachya D.J.Bedford](#); current. Source: Bedford, D.J. (1986)  
[Xanthorrhoea acaulis \(A.T.Lee\) D.J.Bedford](#); new in 1986, current name for [Xanthorrhoea subsp. acaulis A.T.Lee](#). Source: Bedford, D.J. (1986).  
[Xanthorrhoea arborea R.Br.](#); current. Source: Bedford, D.J. (1986).  
[Xanthorrhoea arborea R.Br.](#); now [Xanthorrhoea resinosa Pers.](#). Source: CHAH (2006).  
[Xanthorrhoea arenaria D.J.Bedford](#); current name for [Xanthorrhoea minor R.Br.](#). Source: Bedford, D.J. (1986).

**What's Its Name?** provides nomenclature information on plant names used in Australia.<sup>415</sup> WIN derives its information of names from the Australian Plant Name Index.

- Current/accepted names are presented in roman type
- Synonyms are in *italic* type
- Plant names recognised in the 1990 *Census of Australian Vascular Plants* are accepted unless our review of subsequent literature indicates otherwise. This review is on-going and much remains to be done.
- Occasional comments about the status of the name are based on decisions at the Australian National Herbarium.
- The source of the reference is provided at the end of each line entry.
- Full references are given at the end of the web page.
- Links from plant names take you to the more extensive entry of the *Australian Plant Name Index* (APNI).

<sup>415</sup> <http://www.anbg.gov.au/win/>

While APNI offers a technical presentation of Australia's botanical nomenclature to the scientific community, *What's Its Name (WIN)* was developed as an alternative user-friendly entry point for the non-botanist. *What's Its Name* is derived from the APNI database and aims to deliver a concise overview of plant names and name changes via the Internet. It provides the current name in use for a plant and a cross-reference to any previously used names. In cases of discrepancy or ambiguity WIN provides a simple explanation of the scientific information suitable for the non-scientific user.<sup>416</sup>

#### **Example using *Banksia aquilonia* entry**

Name = *Banksia aquilonia* (A.S.George) A.S.George

Status = new in 1996; current name for *Banksia integrifolia* var. *aquilonia* A.S.George, *Banksia integrifolia* subsp. *aquilonia* (A.S.George) K.R.Thiele

Comment (from Australian National Herbarium) = This recent upgrade from subspecific status has generally been accepted in most Eastern State Herbaria

Information Source (full reference given at end of web page) = Source: George, A.S. (1999)

An alternative non-tabular delivery of the WIN? output can also be generated.

**What's Its Name?** is a collaborative project of the Australian National Botanic Gardens, the Centre for Plant Biodiversity Research, and the Australian Biological Resources Study. It is supported by the Council of Heads of Australian Herbaria.<sup>417</sup>

#### **Function**

- Taxonomy
- User interface
  - Personal use
  - Raw data

#### **Why use this tool?**

- Name information for plants

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens
- Data users
  - Expert
  - Interest groups
- ALA infrastructure

#### **How will the tool be used?**

- Online tool
- User input required

#### **Where in the data chain could this tool be used?**

- Data source
- User's machine

---

<sup>416</sup> <http://www.anbg.gov.au/databases/apni-about/index.html>

<sup>417</sup> <http://www.anbg.gov.au/win/>

**When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request

**Availability**

- WIN? What's Its Name: <http://www.anbg.gov.au/win/>
- ANBG: <http://www.anbg.gov.au/anbg/index.html>
- Cost: free

**Comments**

## Genes

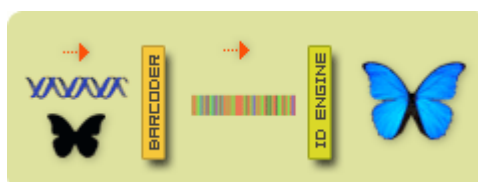
## BOLD-ID Barcode of Life Identification System

Summary	
Type of tool	Application
Function	Species identification using gene sequences
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Operational
Time of use	Data analysis
Licence	Unknown

**Barcode of Life Data Systems (BOLD)** “provides a species identification tool that accepts DNA sequences from the barcode region and returns a taxonomic assignment to the species level when possible”.<sup>418</sup>

### Description

The BOLD Identification System (IDS) accepts sequences from the 5' region of the mitochondrial gene COI and returns a species-level identification when one is possible. Further validation with independent genetic markers will be desirable in some forensic applications.



The reference database of validated records is used by default and is recommended for all identification purposes.<sup>419</sup>

**DNA barcode sequences** are very short relative to the entire genome and they can be obtained reasonably quickly and cheaply. The "Folmer region" at the 5' end of the cytochrome c oxidase subunit 1 mitochondrial region (COI) is emerging as the standard barcode region for almost all groups of higher animals. This region is 648 nucleotide base pairs long in most groups and is flanked by regions of conserved sequences, making it relatively easy to isolate and analyze. A growing number of studies have shown that COI sequence variability is very low (generally less than 1-2%) and that the COI sequences of even closely related species differ by several percent, making it possible to identify species with high confidence. For those groups in which COI is unable to resolve species-level differences, CBOL recommends the use of an additional gene region. In some groups, COI is not an effective barcode region and a different standard region must be identified. In all cases, DNA barcoding is based on the use of a short, standard region that enables cost-effective species identification.<sup>420</sup>

This is part of the BOLD online workbench for assisting with collection, management, analysis and use of DNA barcodes.

BOLD is associated with the following barcoding campaigns:<sup>421</sup>

<sup>418</sup> <http://www.barcodinglife.org/>

<sup>419</sup> <http://www.barcodinglife.org/views/idrequest.php>

<sup>420</sup> <http://barcoding.si.edu/>

<sup>421</sup> <http://www.barcodinglife.org/views/projectlist.php?>



- ACG Parasitoids
- All Birds Barcoding Initiative
- Ants Of The World
- Barcoding Fish (FishBOL)
- Barcoding Mammals of the World
- Fungal Barcoding
- GenBank Animals (COI)
- GenBank Fungi (COI)
- Canadian Barcode of Life Network (BOLNET)
- Consortium for the Barcode of Life (CBOL)
- Lepidoptera
- Mosquitoes of the World
- PolarBOL
- Tephritid Barcode Initiative

#### **Function**

- Identification tools

#### **Why use this tool?**

- To determine the species from an appropriate gene sequence

#### **Who will use this tool?**

- Data creation
  - Expert
- Data users
  - Expert
- DNA skills required

#### **How will the tool be used?**

- Data sequences required in fasta format
- Online tool through web browser
- User input required

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- As a post process, after data is with the user

#### **Availability**

- Identification: <http://www.barcodinglife.org/views/idrequest.php>
- A resource of BOLD: <http://www.barcodinglife.org/>
- Documentation: <http://www.barcodinglife.org/docs/boldmas.html>
- Online tutorial: <http://www.barcodinglife.org/docs/boldtutorial.html>
- Consortium for the Barcode of Life (CBOL): <http://barcoding.si.edu/>
- Barcoding Life - Ten Reasons: <http://barcoding.si.edu/PDF/TenReasonsBarcoding.pdf>
- Cost: free
- Licence unknown

#### **Comments**

BOLD is a resource for the DNA barcode community.<sup>422</sup>

- It supports the organization and analysis of barcode data.
- It provides a repository for barcode records, storing specimen data and images as well as sequences and trace files.
- It provides an efficient interface for submitting barcode records to GenBank.
- It provides an identification engine based on the current barcode library.
- It monitors the number of barcode sequence records and species coverage.

BOLD also encourages external connectivity. They have developed BOLD-ECS which “provides web developers and bioinformaticians the ability to build tools and workflows that can be integrated with the BOLD framework”.<sup>423</sup>

Perhaps it is possible to build into ALA, the ability to directly connect to and integrate with BOLD-ECS.

---

<sup>422</sup> <http://www.barcodinglife.org/>

<sup>423</sup> <http://www.barcodinglife.org/>

## GenBank

Summary	
Type of tool	Database of gene sequences
Function	Taxonomy, gene research
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Active development. Release 163. December 2007
Time of use	At time of user request.
Licence	No restrictions on the use or distribution of the GenBank data

**GenBank** is the NIH genetic sequence database, an annotated collection of all publicly available DNA sequences.<sup>424</sup>

### Description

GenBank depends on its contributors to help keep the database as comprehensive, current, and accurate as possible. NCBI (National Center for Biotechnology Information) provides timely and accurate processing and biological review of new entries and updates to existing entries, and is ready to assist authors who have new data to submit.<sup>425</sup>

```
1 atgaggetcg cttaccttct cctactgctt gtggccgtgt tgttccaagc tgggggcgga
61 tcagccaaac ccataatggt cttcgaatg caagcctgct ggtcccacag cggcgtctgc
121 agagacaaga gtgaaaggaa ttgcaagccg atggcgtgga cctactgtga aaaccggaac
181 cagaaatgct gcgactccgg ccacagctgt cgtgtcatcg tcatgaggct cacttacctt
241 ctctactgct ttgtggccgt gttgttccag gctgggagtg gatcagccga acccatattc
301 ttctatggac gccaacctg ctctactac gacgggtct gccgagacaa aagcgatgtg
361 aattgcaaat atatecgctt cacctactgt gaaaaccgga accagagatg ctgctactac
421 tag
```

mRNA gene sequence of a peptide from platypus venom *Ornithorhynchus anatinus*.<sup>426</sup>

Many journals require submission of sequence information to a database prior to publication so that an accession number may appear in the paper. Software is provided. There are specialized, streamlined procedures for batch submissions of sequences, such as EST, STS, and GSS sequences. Revisions or updates to GenBank entries can be made at any time.<sup>427</sup>

GenBank is available for searching at NCBI via several methods.

GenBank is part of the International Nucleotide Sequence Database Collaboration, which comprises the DNA DataBank of Japan (DDBJ), the European Molecular Biology Laboratory (EMBL), and GenBank at NCBI. These three organizations exchange data on a daily basis.<sup>428</sup>

### Function

- Taxonomy
  - Identification tools
- User interface
  - Personal and institutional
  - Raw data

<sup>424</sup> <http://www.ncbi.nlm.nih.gov/Genbank/>

<sup>425</sup> <http://www.ncbi.nlm.nih.gov/Genbank/submit.html>

<sup>426</sup> [http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?val=XM\\_001519227.1](http://www.ncbi.nlm.nih.gov/entrez/viewer.fcgi?val=XM_001519227.1)

<sup>427</sup> <http://www.ncbi.nlm.nih.gov/Genbank/>

<sup>428</sup> <http://www.ncbi.nlm.nih.gov/Genbank/>

### Why use this tool?

- GenBank should be used to store gene sequence data prior to publication.
- Data in GenBank can be used for taxonomy

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data users
  - Expert

### How will the tool be used?

Tools are available for submission<sup>429</sup>

- Sequin Stand-alone sequence submission tool
- BankIt A web submission tool for quick and simple submissions
- tbl2asn Command-line sequence submission tool

Batch submission is available

GenBank divisions

- **dbEST** accepts Expressed Sequence Tags (ESTs) are short (usually about 300-500 bp), single-pass sequence reads from mRNA (cDNA).<sup>430</sup>
- **dbGSS** is similar in nature to the EST division, except that its sequences are genomic in origin, rather than cDNA (mRNA).<sup>431</sup>
- **dbSTS** accepts short (about 200-500 bp) sequences that are operationally unique in a genome (i.e., can be specifically detected by PCR in the presence of all other genomic sequences), and that define a specific position on the physical map.

Searching

- Entrez Browser GenBank (nucleotides and proteins), PubMed (MEDLINE), 3D structures, genomes, and PopSet databases
- BLAST Sequence Similarity Searching Nucleotide or protein query sequences against the specified database using the BLAST suite of algorithms. GenBank nucleotide records are located in separate databases that must be searched independently.
- Network Entrez is a client/server version<sup>432</sup>

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- At the time of a user request

### Availability

- GenBank: <http://www.ncbi.nlm.nih.gov/Genbank/>
- Entrez search: <http://www.ncbi.nlm.nih.gov/Entrez/>
- PubMed publications: <http://www.ncbi.nlm.nih.gov/sites/entrez?db=pubmed>

---

<sup>429</sup> <http://www.ncbi.nlm.nih.gov/Genbank/submit.html>

<sup>430</sup> [http://www.ncbi.nlm.nih.gov/dbEST/how\\_to\\_submit.html](http://www.ncbi.nlm.nih.gov/dbEST/how_to_submit.html)

<sup>431</sup> [http://www.ncbi.nlm.nih.gov/dbGSS/how\\_to\\_submit.html](http://www.ncbi.nlm.nih.gov/dbGSS/how_to_submit.html)

<sup>432</sup> <http://www.ncbi.nlm.nih.gov/Web/Search/client.html>

- Submission guide: <http://www.ncbi.nlm.nih.gov/Genbank/submit.html>
- **NCBI** National Center for Biotechnology Information: <http://www.ncbi.nlm.nih.gov/>
- Active development. Release 163. December 2007
- Restrictions: no restrictions on the use or distribution of the GenBank data.<sup>433</sup>

### Comments

GenBank accepts *all* DNA sequences. As a consequence of this policy, some of the sequences here are rubbish. Compare this to BOLD which only accepts sequences from the 5' region of the mitochondrial gene COI.

---

<sup>433</sup> <http://www.ncbi.nlm.nih.gov/Genbank/>

## Gene tools

Summary	
Type of tool	Applications
Function	Genetics data analysis
Online / Desktop	Online and desktop
Computer infrastructure	Depends on application
Development status	
Time of use	When data is with user
Licence	Varied

There are many tools used for genetic data analysis. Discussed here are resources where some of these tools can be found.

### Description

Genetics data is quite distinct and many programs have been developed to assist with its analysis. Some programs are multi-purpose, while others are highly specialised. Below is a selection of resources that include links to, and synopses and reviews of, many genetic analysis programs.

1. NCBI Tools for Data Mining (<http://www.ncbi.nlm.nih.gov/Tools/>). Tools have been developed for nucleotide and protein sequence analysis, structures, gene expression, and for genome analysis.
2. An Alphabetic List of Genetic Analysis Software (<http://linkage.rockefeller.edu/soft/list1.html>). Over 400 programs include genetic linkage analysis, QTL analysis for animal/plant breeding data, genetic marker ordering, genetic association analysis, haplotype construction, pedigree drawing, and population genetics.<sup>434</sup>
3. Resources for Ecology, Evolutionary Biology, Systematics, and Conservation Biology (<http://darwin.eeb.uconn.edu/links/category.php?id=12>). Of the many categories of resources here, there are links to molecular evolutionary software.
4. Genetic Data Analysis (<http://gallus.forestry.uga.edu/dataAnalysis/>) The emphasis is on analysis/generation of genetic data using maximum likelihood or Bayesian estimation of parameters of interest.<sup>435</sup>
5. The Gene Ontology - Tools for Gene Expression Analysis (<http://www.geneontology.org/GO.tools.microarray.shtml>). Software developed to use the gene ontologies or gene associations provided by GO consortium members.
6. GSF - The Genetic Software Forum (<http://www.rannala.org/gsf/>). Discussion lists, Q&A, and articles about genetic software
7. A recent review paper which describes the principles, the statistics computed, the assumptions made, and some limitations, is: Excoffier L & Heckel G (2006) *Computer programs for population genetics data analysis: a survival guide*. **Nature Reviews Genetics** 7: 745-758, (October 2006).

<sup>434</sup> <http://linkage.rockefeller.edu/soft/list1.html>

<sup>435</sup> <http://gallus.forestry.uga.edu/dataAnalysis/>

## Function

- Analysis tools

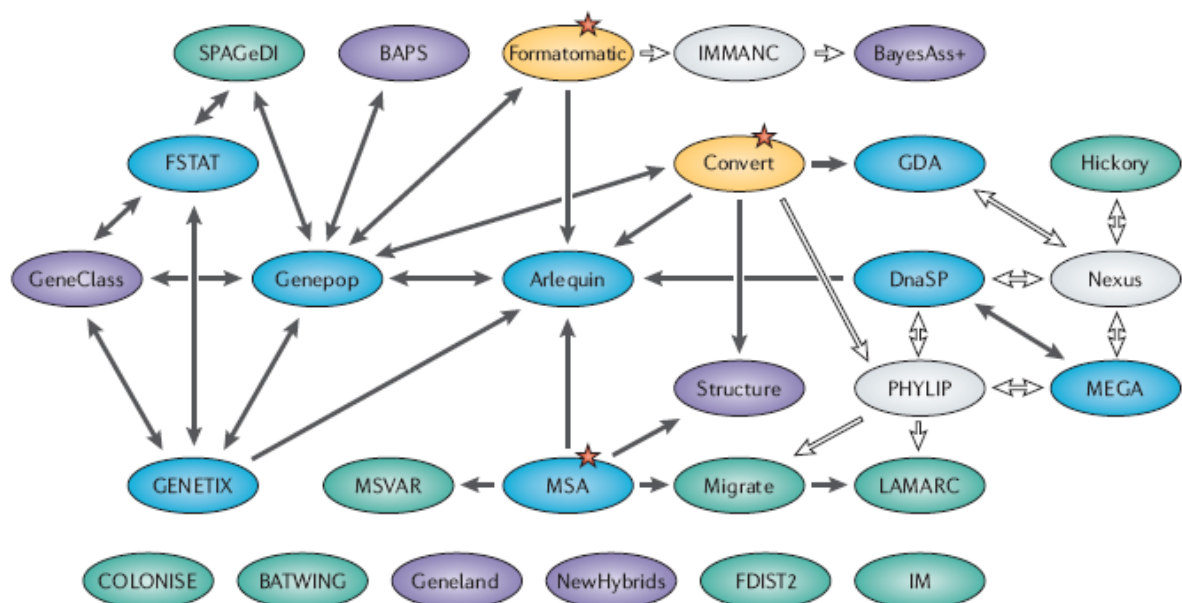
## Why use this tool?

- Genetics analyses

## Who will use this tool?

- Data users
  - Expert
  - Interest groups
- Special skills are required

## How will the tool be used?



Flow chart of possible data exchange between different population genetics programs. Although many programs have their own input-file specification, data files can still be exchanged between most programs (black arrows), avoiding tedious reformatting processes. The red stars are recommended starting points to format an initial data set. Blue ellipses represent multi-purpose packages, whereas individual-centred programs are shown in violet. The two conversion programs are shown in yellow. Specialized programs are shown in green, and light grey ellipses represent programs ... the data formats of which are used by other programs allowing indirect data exchange (white arrows). The data files associated with the programs listed on the bottom row cannot be exchanged directly with the other programs.<sup>436</sup>

- Online and desktop applications
- Available for many different platforms
- User input is required

## Where in the data chain could this tool be used?

<sup>436</sup> Excoffier L & Heckel G, *Computer programs for population genetics data analysis: a survival guide*. **Nature Reviews Genetics** 7: 745-758 (October 2006)  
[abstract: <http://www.nature.com/nrg/journal/v7/n10/abs/nrg1904.html>]

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- NCBI Tools for Data Mining: <http://www.ncbi.nlm.nih.gov/Tools/>
- Alphabetic List of Genetic Analysis Software: <http://linkage.rockefeller.edu/soft/list1.html>
- Resources for Ecology, Evolutionary Biology, Systematics, and Conservation Biology: <http://darwin.eeb.uconn.edu/links/category.php?id=12>
- Genetic Data Analysis: <http://gallus.forestry.uga.edu/dataAnalysis/>
- The Gene Ontology: <http://www.geneontology.org/GO.tools.microarray.shtml>
- GSF - The Genetic Software Forum: <http://www.rannala.org/gsf/>
- Many genetics analysis programs are freely available

**Comments**

- Rather than manipulate genetics data specifically, the ALA could probably act as a conduit or link between the source and the user.
- This list of resources is by no means complete.



## Swami - The Next Generation Biology Workbench

Summary	
Type of tool	Framework with tools
Function	Gene data manipulation
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Active development, in draft release. Version 0.9, October 2007
Time of use	As a post process, when data is with user
Licence	Free for research and education

**The Next Generation Biology Workbench** is a set of tools for gene data manipulation and analysis

### Description

The Next Generation Biology Workbench is a free resource for research and education in Bioinformatics, Genomics, Proteomics, and Phylogenetics.

This is a resource that is easy to use, and can accomplish many different kinds of tasks. The toolkit is small today, but will grow quickly. We expect to release between 100 and 200 tools in December (2007), along with new databases and improved data and tool handling semantics.<sup>437</sup>

### Tools

- BLAST (v2.214)
- Boxshade (v3.3.1)
- CLIQUE (v3.682)
- ClustalW (v1.82)
- CONSENSE (v3.682)
- DNADIST (v3.682)
- DNAPARS (v3.682)
- DOLLOP (v3.682)
- MFOLD (v3.1.2)
- MIX (v3.682)
- PARS (v3.682)
- PROTDIST (v3.682)
- PROTPARS (v3.682)
- PSIBLAST (v2.2.1)
- SEQBOOT (v3.682)
- UNROOT (v3.682)
- SIRIUSWB

Swami will support protein structure investigations and allow file sharing.<sup>438</sup>  
Discrete analytical tasks can be chained.

### Function

- Analysis tools
- User interface

<sup>437</sup> <http://www.ngbw.org/web/welcome.action>

<sup>438</sup> "The Next Generation Biology Workbench" (2007) Mark A. Miller ISMB2007; Vienna, Austria.  
[http://www.ngbw.org/pages/static/publications/ISMB\\_2007\\_7\\_24\\_workbench.ppt](http://www.ngbw.org/pages/static/publications/ISMB_2007_7_24_workbench.ppt)

- Personal use
- Can produce reports, web pages, further data sets

**Why use this tool?**

- Gene sequence analysis

**Who will use this tool?**

- Data users
  - Expert

**How will the tool be used?**

- Online tool, accessible through most browsers<sup>439</sup>
- User input is required

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- Swami: <http://www.ngbw.org/web/welcome.action>
- Feedback: <http://www.ngbw.org/web/contact.action>
- Draft release. Version 0.9 October 2007. Beta version due December 2007
- Users need to register, a guest account is available
- Cost: Free

**Comments**

---

<sup>439</sup> <http://www.ngbw.org/web/compatibility.action>

## Metadata

## ASDD – Australian Spatial Data Directory

Summary	
Type of tool	Directory
Function	Metadata directory
Online / Desktop	Online
Computer infrastructure	Z39.50 connection
Development status	Active. Over 20 participants
Time of use	While data is stored with ALA, at time of user request
Licence	

**ASDD – Australian Spatial Data Directory** is a spatial metadata search facility for the disparate Australian spatial data providers.

### Description

The ASDD is an essential component of the Australian Spatial Data Infrastructure (ASDI). The gateway to the ASDD is maintained by Geoscience Australia on behalf of ANZLIC - the Spatial Information Council, as part of its broader Australian Government responsibility for the ASDI. The individual ASDD nodes are implemented by State/Territory jurisdictions, Australian Government agencies, and commercial organisations (for details see all node descriptions). It is the responsibility of individual nodes to maintain their own metadata and nodes in accordance with the ANZLIC Metadata Guidelines and the ASDD Requirements and Standards.<sup>440</sup>

The Australian Spatial Data Directory (ASDD) provides these search interfaces to discover geospatial dataset descriptions (metadata) throughout Australia. A dataset description is a concise document which consistently explains a certain set of geospatial (earth-related) data, and provides links to further information and possibly to the actual data.

The ASDD is an ANZLIC - the Spatial Information Council Australian Spatial Data Infrastructure (ASDI) initiative.

OSDM is sponsoring a pilot project to replace the ASDD gateways with GeoNetwork. More information about this can be found at <http://asdd.ga.gov.au/asdd/asdi/>.<sup>441</sup>

### About the ASDD<sup>442</sup>

The Australian Spatial Data Directory (ASDD) is a national initiative supported by all governments under the auspices of ANZLIC - the Spatial Information Council. The ASDD aims to improve access to Australian spatial data for industry, government, education and the general community through effective documentation, advertisement and distribution. The directory comprises government and commercial nodes in each State/Territory and spatial data agencies within the Australian Government.

A key objective of the ANZLIC strategic plan is to promote the development of the Australian Spatial Data Infrastructure (ASDI) which will improve access to and availability of nationally consistent spatial datasets. The ASDD is an essential component of the ASDI and incorporates information about datasets (metadata) from all jurisdictions.

<sup>440</sup> <http://asdd.ga.gov.au/>

<sup>441</sup> <http://asdd.ga.gov.au/>

<sup>442</sup> <http://asdd.ga.gov.au/asdd/about.htm>

The ASDD was launched in 1998 and has since steadily grown in content to become the key source of spatial information in Australia.

The technology being used for the ASDD is the Z39.50 search and retrieval protocol which when combined with the World Wide Web provides a simple method of searching, discovery and retrieval of spatial data. More detail is available in the Technical Documentation which also includes instructions for establishing a node of the ASDD.

Geoscience Australia maintains the gateway on behalf of ANZLIC, as part of its broader Australian Government responsibility for the Australian Spatial Data Infrastructure.

#### **Becoming a node of the ASDD<sup>443</sup>**

To participate in the ASDD, an organisation needs to meet certain criteria. These are summarised in Requirements and Standards.

#### **ASDD Nodes<sup>444</sup>**

- [ACT Geographic Data Directory](#)
- [Australian Antarctic Data Centre](#)
- [Australian Hydrographic Service - Product Metadata](#)
- [Australian Hydrographic Service - Publication Metadata](#)
- [Australian Hydrographic Service - Source Metadata](#)
- [BRS and Australian Natural Resources Data Library \(ANRDL\)](#)
- [Bureau of Meteorology](#)
- [CSIRO Marine and Atmospheric Research](#)
- [DEW Discover Information Geographically \(DIG\)](#)
- [Geoscience Australia](#)
- [IndexGeo Pty Ltd - Eco Companion catalogue](#)
- [Murray-Darling Basin Commission](#)
- [National Oceans Office](#)
- [NSW Natural Resources Data Directory](#)
- [NT Spatial Data Directory](#)
- [Other Commonwealth Agencies \(hosted by BRS\)](#)
- [PSMA Australia Limited](#)
- [Qld Spatial Data \(QSIC\)](#)
- [Queensland Department of Natural Resources and Mines Spatial Data](#)
- [RAN Directorate of Oceanography and Meteorology](#)
- [South Australian Spatial Information Directory](#)
- [Tasmanian Spatial Data Directory](#)
- [Victorian Spatial Data Directory](#)
- [WALIS Interragator-Aerial Photography](#)
- [WALIS Interragator-Bibliographic Data](#)
- [WALIS Interragator-Environmental Impact Statements](#)
- [WALIS Interragator-Natural Resource Monitoring](#)
- [WALIS Interragator-Spatial Data](#)

---

<sup>443</sup> <http://asdd.ga.gov.au/asdd/about.htm>

<sup>444</sup> <http://asdd.ga.gov.au/asdd/tech/node/>

An example of node metadata<sup>445</sup>

**Bureau of Meteorology**

**Node type**  
Commonwealth government

**Node description**  
Selected Bureau of Meteorology base data sets.

**Z39.50 connection details**  
True Z39.50 client software can also be used to search this collection.  
See lists at [Z39.50 Maintenance Agency](#)

- **Server hostname:** extapps.bom.gov.au
- **Server port:** 6668
- **Repository name:** bom\_meta

**Bounding coordinates**  
The outer geographic extent of the whole document collection ...

- **North bounding coordinate:** -9
- **South bounding coordinate:** -44
- **East bounding coordinate:** 154
- **West bounding coordinate:** 112

**Top search terms**  
These terms occur frequently in the metadata collection. They will give you a starting point for conducting searches.  
Forecast, Rainfall, District, Boundary, Meteorology, Climate

**References**

- **Home:** <http://www.bom.gov.au/>

**Contact details**

- **Contact organisation:** Bureau of Meteorology
- **Postal address:** PO box 1289K
- **City:** Melbourne
- **State:** Victoria
- **Postal code:** 3001
- **Electronic mail:** mgdu@bom.gov.au
- **Telephone:** 0396694474

**Metadata date**  
2004-05-11

The ASDD checks the status of each of the connected nodes every hour. Number of records available and response time is logged. A summary report is generated monthly and made publicly accessible.

Total ASDD records 40,000 (December 2007).

---

<sup>445</sup> <http://asdd.ga.gov.au/asdd/tech/node/bom-1.html>

## Function

- Metadata
  - Data discovery
- User interface
  - Institutional use
  - Raw data

## Why use this tool?

- Access to Australian spatial metadata

## Who will use this tool?

- ALA infrastructure

## How will the tool be used?

- Z39.50 search and retrieval protocol
- Runs as a batch job

## ASDD Metadata Search

### Basic Search<sup>446</sup>

- AND, OR, or PHRASE

### Advanced Search<sup>447</sup>

- Boolean searches
- HTML, XML, plain text
- Node selection
- Date selection
- Keyword search
- Spatial selection – geographic coordinates, define a region using a map, or select from a pre-defined list of geographic extents (eg. State or Territory)

## Where in the data chain could this tool be used?

- ALA central

## When could this tool be used?

- While data is stored with ALA
- Perhaps at the time of a user request

## Availability

- ASDD: <http://asdd.ga.gov.au/>
- ANZLIC: <http://www.anzlic.org.au/>
- Documentation: <http://asdd.ga.gov.au/asdd/tech/>
- Cost: Free access to metadata.
- Restrictions: Many databases are restricted.

## Comments

---

<sup>446</sup> <http://asdd.ga.gov.au/asdd/tech/zap/basic.html>

<sup>447</sup> <http://asdd.ga.gov.au/asdd/tech/zap/advanced.html>

## DIG – Discover Information Geographically

Summary	
Type of tool	Directory
Function	Data discovery
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Ailing
Time of use	At time of user request
Licence	

**Discover Information Geographically (DIG)** is a tool which provides information regarding whether the Department of the Environment, Water, Heritage and the Arts holds data on a particular topic of interest or for an area of interest. Entries currently cover data and publications, *but will be extended to incorporate map services and printed maps.*<sup>448</sup>

### Description

All DIG entries are related to a geographical location, and most of the data referred to is suitable for use in a GIS (Geographic Information System). Data is collected at a national (rather than regional) scale, and covers both the continent and international waters.

Only data which is maintained by the Department, or may be supplied to third parties, is available. Supply of any of the data sets will involve the acceptance of a licence.

DIG is updated as soon as new, updated or replacement data sets become available for general use. In some instances, several versions of a data set may be available, because the previous versions are of use for comparison purposes.

### How to use DIG

A search is conducted using keywords, categories or by defining a geographical area. A brief overview of each entry matching the criteria is displayed with links to the full entry and a simple map displaying the approximate geographic location of the data.

Alternatively, use the BROWSE tab to browse all metadata entries, which are organised in topic folders.

The referenced data may not be available on this web site; however, relevant contact details are included with every entry. Where data sets are available for download, licence details are included.<sup>449</sup>

**Content** includes:<sup>450</sup>

- Data
  - Live data and maps
  - Downloadable data
  - Departmental data
  - Archived data
- Documents

<sup>448</sup> <http://www.environment.gov.au/erin/dig/>

<sup>449</sup> <http://www.environment.gov.au/erin/dig/>

<sup>450</sup> <http://www.environment.gov.au/metadateexplorer/explorer.jsp>



- Map files
- Static map images
- Resources
  - Applications
  - Geographic services
  - Geographic activities
  - Clearing houses

### **Categories**<sup>451</sup>

- Agriculture
- Base maps
- Biota
- Boundaries
- Cadastre
- Climatology
- Economy
- Elevation
- Environment
- Geoscience
- Health
- Inland Waters
- Location
- Military
- Oceans
- Society
- Structures
- Transportation
- Utilities

### **Function**

- Metadata
  - Data discovery
- User interface
  - Personal use
  - Visual presentation

### **Why use this tool?**

- To discover whether the Department of the Environment, Water, Heritage and the Arts holds data sets of interest

### **Who will use this tool?**

- Data users
  - Expert
- ALA infrastructure

### **How will the tool be used?**

- Online search
- Browser

---

<sup>451</sup> <http://www.environment.gov.au/metadataexplorer/explorer.jsp>

- User input required
- Can it run as a batch job?

**Where in the data chain could this tool be used?**

- ALA central
- User's machine

**When could this tool be used?**

- While data is stored with ALA
- At the time of a user request

**Availability**

- DIG: <http://www.environment.gov.au/erin/dig/>
- Search: <http://www.environment.gov.au/metadataexplorer/explorer.jsp>
- Restrictions: The referenced data may not be available on this website, however, relevant contact details are included with every entry. Where data sets are available for download, licence details are included.<sup>452</sup>

**Comments**

Can be very slow (more than 1 minute to return search result), and prone to errors.

---

<sup>452</sup> <http://www.environment.gov.au/erin/dig/index.html>

## docBUILDER

Summary	
Type of tool	Application
Function	Metadata creation
Online / Desktop	Online
Computer infrastructure	Data held by Global Change Master Directory (GCMD)
Development status	Last modified November 2007
Time of use	Data creation, data capture
Licence	

### Description

The **docBUILDER** tool is designed for metadata authors to add (or modify) data set descriptions (DIFs), related services descriptions (SERFs) and descriptions of Ancillary Information (Data Centres, Projects, Instruments and Platforms).<sup>453</sup>

### Function

- Data Identification tool

### Why use this tool?

Use docBUILDER to create a new metadata record.

### Who will use this tool?

- Data creators
  - Experts, scientists
- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
  - Casual users

### How will the tool be used?

- This is a web based tool used online. The metadata remains at GCMD
- User input required
- Templates can be created for creating multiple records with common factors

### Where in the data chain could this tool be used?

- Data source

### When could this tool be used?

- Before data is made available to ALA

### Availability

docBUILDER and Master Directory software are produced in collaboration with NASA's Goddard Space Flight Center through the Global Change Master Directory (GCMD) project.<sup>454</sup>

- Online location: <http://gcmd.nasa.gov/User/authoring.html>

---

<sup>453</sup> <http://gcmd.nasa.gov/DocumentBuilder/Home.do>

<sup>454</sup> <http://gcmd.nasa.gov/DocumentBuilder/Home.do>

- Online guide: <http://gcmd.gsfc.nasa.gov/DocumentBuilder/Home.do?RequestAction=Help>
- GCMD - Global Change Master Directory: <http://gcmd.nasa.gov/>
- Last updated November 2007
- Privacy and security notices: [http://www.nasa.gov/about/highlights/HP\\_Privacy.html](http://www.nasa.gov/about/highlights/HP_Privacy.html)

### Comments

This service is built for the earth science community and is cross mapped to other standards such as the National Biological Information Infrastructure (NBII), the Dublin Core Metadata Initiative, and the Australia New Zealand Land Information Council (ANZLIC).

This metadata tool is used by the Australian Antarctic Data Centre, and metadata records are entered directly into the Antarctic Master Directory maintained by the [Global Change Master Directory \(GCMD\)](#).

Ocean Biogeographic Information System (OBIS) have a plan to use GCMD docBUILDER and metadata facilities.<sup>455</sup>

Portals are available for partner organisations wishing to customise the view of the GCMD directory.

---

<sup>455</sup> <http://www.iobis.org/tech/metadata1/gcmd-plan> (MS-Word document)

## EML - Ecological Metadata Language

Summary	
Type of tool	Language
Function	Metadata language
Online / Desktop	Online validation service
Computer infrastructure	
Development status	version 2.0.1 August 2004
Time of use	Data creation
Licence	Open source

The **EML** project is an open source, community oriented project dedicated to providing a high-quality metadata specification for describing data relevant to the ecological discipline.<sup>456</sup>

### Description

Ecological Metadata Language (EML) is a metadata specification developed by the ecology discipline and for the ecology discipline.

EML is implemented as a series of XML document types that can be used in a modular and extensible manner to document ecological data. Each EML module is designed to describe one logical part of the total metadata that should be included with any ecological dataset.<sup>457</sup>

### Function

- Identification tools

### Why use this tool?

With this tool the ecological community has an extensible, flexible, metadata standard for use in data analysis and archiving that allows automated machine processing, searching and retrieval.<sup>458</sup>

### Who will use this tool?

- Data creation
  - Experts
- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
  - Casual users
- ALA infrastructure

### How will the tool be used?

- Specification in the form of XML schema
- Validation tool available online
- User input required for validation

---

<sup>456</sup> <http://knb.ecoinformatics.org/software/eml/>

<sup>457</sup> <http://knb.ecoinformatics.org/software/eml/>

<sup>458</sup> EML 2.0.1 Specification

**Where in the data chain could this tool be used?**

- Data source
- ALA central could use the specification

**When could this tool be used?**

- Before data is made available to ALA

**Availability**

- Project: <http://knb.ecoinformatics.org/software/eml/>
- Download: <http://knb.ecoinformatics.org/software/download.html#eml>
- Source code: <http://cvs.ecoinformatics.org/cvs/cvsweb.cgi/eml/>
- FAQ: <http://knb.ecoinformatics.org/software/eml/eml-2.0.1/eml-faq.html>
- Specification: <http://knb.ecoinformatics.org/software/eml/eml-2.0.1/index.html>
- Version: EML 2.0.1 August 2004
- Licence: Open Source under the terms of the GNU GPL  
<http://knb.ecoinformatics.org/software/license.html>
- Free to download
- Validation service: <http://knb.ecoinformatics.org/emlparser/> confirms EML file conforms to the EML specification.

**Comments**

This is a language specification.

## MMI – Marine Metadata Interoperability

Summary	
Type of tool	Collaborative portal with tools
Function	Metadata services
Online / Desktop	Online and desktop
Computer infrastructure	Various
Development status	Fresh
Time of use	Before data is made available to ALA
Licence	

**MMI – Marine Metadata Interoperability** is developing web applications and stand-alone tools to enable sophisticated interactions across marine data systems.

### Description

MMI creates tools to make the data management process more efficient. Using MMI tools, data technologists can increase the interoperability of the scientific data they manage, providing additional benefits to scientific resource providers and users.<sup>459</sup>

MMI describes three classes of technical metadata standards – transport protocols, content standards, and vocabularies – and provides pointers to key tools that are used to work with metadata. MMI collects and develops guides on data management approaches.<sup>460</sup>

Tools developed at MMI for metadata:<sup>461</sup>

- **Vocabulary Integration Environment (VINE)** - a tool to map vocabularies.
- **Vocabularies to OWL (voc2owl) Tool** - a tool to perform fast conversion from vocabularies found in plain ASCII format to the Web Ontology Language (OWL).

MMI is funded in part by the National Science Foundation with support and contributions from the marine community.

### Function

- Provider interaction
  - Data preparation
- Metadata
  - Creation
  - Data discovery
- User interface
  - institutional use
  - Raw data

### Why use this tool?

To encourage the discovery and re-use of scientific data.

---

<sup>459</sup> <http://marinemetadata.org/for-scientists>

<sup>460</sup> <http://marinemetadata.org/fordevelopers>

<sup>461</sup> <http://marinemetadata.org/examples/mmihostedwork/tools/>

If a data set is well documented and it is available to other researchers, the data is more likely to be used. Both research efforts will be complemented, opportunities for collaboration will increase, and the scope of research will broaden.<sup>462</sup>

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users
- ALA infrastructure
- Special skills are required

#### **How will the tool be used?**

- Tool dependent

#### **Where in the data chain could this tool be used?**

- Data source
- ALA central

#### **When could this tool be used?**

- Before data is made available to ALA
- While data is stored with ALA

#### **Availability**

- MMI: <http://marinemetadata.org/>
- Metadata tools for creation, managing and access: <http://marinemetadata.org/fordevelopers/tools>
- MMI guides: <http://marinemetadata.org/guides>

#### **Comments**

MMI appear to be a recent arrival with a strong emphasis on promoting the exchange and use of marine data. Although there is not a lot of depth here now, this could be an important group to watch as tools and standards are created, and expectations develop into products.

---

<sup>462</sup> <http://marinemetadata.org/for-scientists>



## M3Cat

Summary	
Type of tool	Application
Function	Metadata Cataloguing tool
Online / Desktop	Online
Computer infrastructure	Windows
Development status	Version 1.5
Time of use	Data creation
Licence	

**M3Cat** is a Multistandard, Multilingual, Metadata Cataloguing Tool for the creation of geospatial metadata.

### Description

M<sup>3</sup>CAT is a tool that assist users in entering and managing metadata about geospatial data sets. Online systems for handling metadata need to rely on their being predictable in both form and content. Predictability is assured only by conformance to standards. There are many standards for metadata such as the FGDC Content Standard for Digital Geospatial Metadata, the GILS standard, the NBII standard (<http://www.nbii.gov/>) and the ISO/TC211 metadata standard. Organisations also tend to define a profile of metadata for their own applications, encompassing validation rules and customisation which take into account their particularities.

Developed under the CGDI Geoinnovations program by Inteltec Geomatics (Inteltec) and its partners.<sup>463</sup>

### Function

- Identification tools

### Why use this tool?

- M3Cat allows users to configure their own metadata fields using the standards delivered with the tool, such as FGDC, GILS, NBII and ISO 19115. Alternatively, users can build a specific standard profile, including language and cultural particularities, to fulfil their needs;
- it allows users to customize the interface for standard values (pull-down lists, etc.) or establish default values for numerous metadata fields (organisation name, projection system, etc.);
- it allows users to use templates that correspond to the type of data set they document and then take advantage of cataloguing aids such as Wizards or metadata inheritance between data sets to facilitate the development of metadata databases;
- it uses an integrated multi-projection, multi-datum geographic viewer to enter data set bounding coordinates.<sup>464</sup>

### Who will use this tool?

- Data creation
  - Experts
  - Skilled
- Data capture

<sup>463</sup> [http://www.geoconnections.org/projects/geoinnovations/1999/Mcat3/m3cat-promo-en.html#BM3\\_3](http://www.geoconnections.org/projects/geoinnovations/1999/Mcat3/m3cat-promo-en.html#BM3_3)

<sup>464</sup> <http://www.intelec.ca/html/en/technologies/m3cat.html>

- Curators – specimens, identification
- Data providers
  - Institutions
  - Private collections
  - Casual users

**How will the tool be used?**

- M3Cat is used within a browser
- Installation of M3Cat requires Internet Information Server under Windows
- Metadata elements stored in either Access or Oracle
- M3Cat can be used in standalone or multiple users mode.

**Where in the data chain could this tool be used?**

- Data source

**When could this tool be used?**

- Before data is made available to ALA

**Availability**

- Download: <http://www.intelec.ca/html/en/telechargement/index.html>
- Support: Intelec <http://www.intelec.ca>
- Cost: free
- Version 1.5

**Comments**

## **Non-biological data**

## ANRDL – Australian Natural Resources Data Library

Summary	
Type of tool	Application
Function	Natural resources data
Online / Desktop	Online
Computer infrastructure	Browser, XML
Development status	Stale
Time of use	At the time of a user request
Licence	



ANRDL gives online access to data held by Australian Natural Resources Atlas

### Description

The **Australian Natural Resources Data Library** is managed by the [Bureau of Rural Sciences](#) and provides a search facility to download data presented in the Australian Natural Resources Atlas, as well as many other data and information related to Australia's natural resources.

The results will take you to the metadata report.<sup>465</sup> Some data may be available for download.

### Function

- Metadata
  - Data discovery
- Non-biological data
  - Environmental data
- User interface
  - Personal use
  - Raw data

### Why use this tool?

- To access data in the Australian Natural Resources Atlas

### Who will use this tool?

- Data users
  - Expert
  - Interest groups
- Some skills required

### How will the tool be used?

- Online
- Browser
- Metadata is available as a visual display or XML
- User input required

### Where in the data chain could this tool be used?

- User's machine

---

<sup>465</sup> <http://www.anra.gov.au/data.html>

**When could this tool be used?**

- At the time of a user request

**Availability**

- Australian Natural Resources Data Library: <http://adl.brs.gov.au/anrdl/php/>
- Australian Natural Resources Atlas: <http://www.anra.gov.au/>
- Australia's Resources Online: <http://www.anra.gov.au/aro/>
- Map Maker: <http://www.anra.gov.au/mapmaker/mapservlet?app=anra>

**Comments**

- This is a troublesome site to use for first time users. Some examples would help a lot.
- There is apparently a lot of potentially useful data here that is well hidden. A browse function would be useful.
- Once the metadata is downloaded, it is not clear whether access to the data is possible

## DEM – Digital Elevation Model

Summary	
Type of tool	Dataset
Function	Topographic elevation
Online / Desktop	Desktop
Computer infrastructure	Format: GRIDASCII, ASCII XYZ,
Development status	Developed
Time of use	Before data is sent to ALA, at ALA, at time of user request
Licence	Use reproduce, adapt and distribute data and derivatives to end-users

A **Digital Elevation Model (DEM)** is a nation-wide regular grid of terrain elevation.

### Description

A gridded DEM is one in which the elevation points are spaced at regular intervals so as to create a grid or lattice. These grids can be directly observed, however, generally they are computed from other elevation information such as contours or irregularly spaced spot heights.<sup>466</sup>

DEMs are available for a grid spacing of 18 seconds, 9 seconds and 3 seconds in longitude and latitude (approximately 500, 250 and 90 metres).

The *9 Second DEM Version 2* has been developed with the objective to maximise usefulness of the product for a wide range of applications. For example, the elevation of source data high points (hills or mountains) is well represented in the *9 Second DEM Version 2*.<sup>467</sup>

The 1:250 000 source scale of the elevation grid makes the product useful for national, State-wide and regional applications.

### Specifications

#### **3 second grid**

Coverage: Selected parts of Australia

Coordinates: Geographical

Datum: AGD66, AHD

#### **9 second grid**

Coverage: Australia

Coordinates: Geographical

Datum: GDA94, AHD

#### **18 second grid**

Coverage: selected parts of Australia

Coordinates: Geographical

Datum: AGD66, AHD

### Function

- Non-biological data
  - Environmental data

<sup>466</sup> <http://www.ga.gov.au/nmd/products/digidat/dem.htm>

<sup>467</sup> [http://www.ga.gov.au/nmd/products/digidat/dem\\_9s.htm](http://www.ga.gov.au/nmd/products/digidat/dem_9s.htm)

- User interface
  - institutional use
  - Raw data

### Why use this tool?

The *9 Second DEM* is particularly useful where accurate representations of surface shape and drainage structure are required. It includes point elevation data, elevation contours, stream lines and cliff lines. The product can also be used for applications requiring accurate representation of absolute elevation values.<sup>468</sup>

### Who will use this tool?

- Data capture
  - Curators
- Data providers
  - Institutions
  - Private collections
- Data users
  - Expert
  - Interest groups
- ALA infrastructure

### How will the tool be used?

- Format: GRIDASCII, ASCII XYZ, ERS
- CD and free download for 9 second DEM
- Desktop use

### Where in the data chain could this tool be used?

- Data source
- ALA central

### When could this tool be used?

- Before data is made available to ALA
- While data is stored with ALA
- At the time of a user request

### Availability

- Geoscience Australia DEM: <http://www.ga.gov.au/nmd/products/digidat/dem.htm>
- Cost: free download 9 second DEM; \$99 CD for 3 second, 9 second, 18 second<sup>469</sup>
- Copyright Commonwealth of Australia
- Licence: Use reproduce, adapt and distribute data and derivatives to end-users<sup>470</sup>  
<http://www.ga.gov.au/nmd/products/purchasing/licencing.htm>

### Comments

<sup>468</sup> [http://www.ga.gov.au/nmd/products/digidat/dem\\_9s.htm](http://www.ga.gov.au/nmd/products/digidat/dem_9s.htm)

<sup>469</sup> <http://www.ga.gov.au/nmd/products/purchasing/licencefees.jsp>

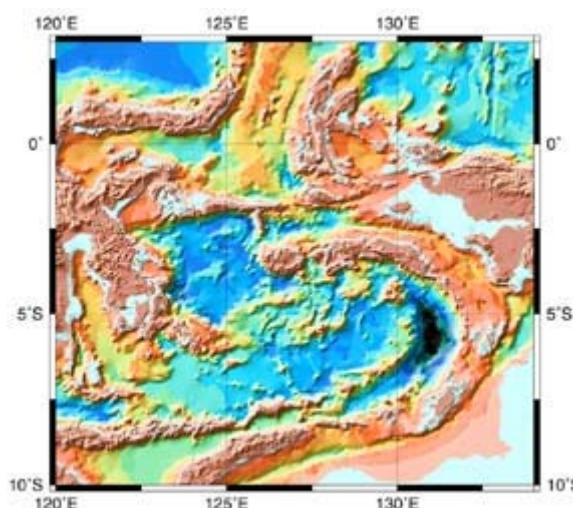
<sup>470</sup> [http://www.ga.gov.au/image\\_cache/GA2716.pdf](http://www.ga.gov.au/image_cache/GA2716.pdf) and [http://www.ga.gov.au/image\\_cache/GA2717.pdf](http://www.ga.gov.au/image_cache/GA2717.pdf)

## GEBCO – General Bathymetric Chart of the Oceans

Summary	
Type of tool	Database
Function	Bathymetric data
Online / Desktop	Desktop
Computer infrastructure	netCDF files, GIS
Development status	CD released April 2003, Bathymetric data
Time of use	At the time of a user request
Licence	Non-commercial reproduction in derivative form with accreditation

The GEBCO Digital Atlas (GDA) consists of a global set of digital bathymetric contours and a one minute interval grid.<sup>471</sup>

### Description



Sample produced from the GEBCO One Minute Grid.<sup>472</sup>

The GDA global set of digital bathymetric contours includes coastlines, trackline control information and a global one arc-minute bathymetric grid — the **GEBCO One Minute Grid**.<sup>473</sup> It also contains the gazetteer of undersea feature names and coastline databases.<sup>474</sup> The complete GDA is distributed on CDROM and is accompanied by full documentation and interface software.<sup>475</sup>

The **GEBCO One Minute Grid** provides bathymetry data on a global grid with a one arc-minute spacing and is available as a download from GEBCO in the form of netCDF files.<sup>476</sup>

Available for free download:

- Data for a user-defined area
- The complete global grid file
- Data in the form of 20 degree square tiles

<sup>471</sup> [http://www.bodc.ac.uk/data/online\\_delivery/gebco/](http://www.bodc.ac.uk/data/online_delivery/gebco/)

<sup>472</sup> <http://www.ngdc.noaa.gov/mgg/gebco/grid/1mingrid.html>

<sup>473</sup> [http://www.bodc.ac.uk/projects/international/gebco/gebco\\_digital\\_atlas/](http://www.bodc.ac.uk/projects/international/gebco/gebco_digital_atlas/)

<sup>474</sup> <http://www.ngdc.noaa.gov/mgg/gebco/gebcocentenarycdrom.html>

<sup>475</sup> [http://www.bodc.ac.uk/projects/international/gebco/gebco\\_digital\\_atlas/](http://www.bodc.ac.uk/projects/international/gebco/gebco_digital_atlas/)

<sup>476</sup> [http://www.bodc.ac.uk/data/online\\_delivery/gebco/](http://www.bodc.ac.uk/data/online_delivery/gebco/)



The grid is based on the most recent version of the GDA bathymetric contours. Additional control contours and sounding data were used in many regions, particularly shallow water areas and semi-enclosed seas, to constrain the gridding process. It is a **continuous digital terrain model for ocean and land**, with land elevations derived from the Global Land One-km Base Elevation (GLOBE) database.

It must be stressed that although the GEBCO One Minute Grid is presented at one minute intervals of latitude and longitude, this does not imply that knowledge is available on sea floor depth at this resolution. It is important to note that, in most places, many miles exist between adjacent ship tracklines and that the grid is an interpolation based upon the input data.<sup>477</sup>

GEBCO's aim is to provide the most authoritative, publicly-available bathymetry for the world's oceans. It operates under the joint auspices of the Intergovernmental Oceanographic Commission (IOC) and the International Hydrographic Organization (IHO).

### Function

- Visualisation tools
  - Maps
- Non-biological data
  - Environmental data
- User interface
  - institutional use
  - Raw data

### Why use this tool?

- Digital bathymetric data globally, including Australian waters

### Who will use this tool?

- Data providers
  - Institutions
- Data users
  - Expert
  - Interest groups
- GIS skills required

### How will the tool be used?

- GEBCO one minute grid is available as netCDF files<sup>478</sup>
- netCDF files are described at World Ocean Circulation Experiment (WOCE) netCDF format<sup>479</sup>
- The GEBCO Digital Atlas is available as a CD, complete with its own interface software<sup>480</sup>
- Data can be exported for GIS compatibility
- Desktop use

### Where in the data chain could this tool be used?

- ALA central

---

<sup>477</sup> [http://www.bodc.ac.uk/data/online\\_delivery/gebco/](http://www.bodc.ac.uk/data/online_delivery/gebco/)

<sup>478</sup> [http://www.bodc.ac.uk/data/online\\_delivery/gebco/](http://www.bodc.ac.uk/data/online_delivery/gebco/)

<sup>479</sup> [http://www.bodc.ac.uk/data/online\\_delivery/international\\_sea\\_level/woce\\_netcdf.html](http://www.bodc.ac.uk/data/online_delivery/international_sea_level/woce_netcdf.html)

<sup>480</sup> [http://www.bodc.ac.uk/projects/international/gebco/gebco\\_digital\\_atlas/](http://www.bodc.ac.uk/projects/international/gebco/gebco_digital_atlas/)

### When could this tool be used?

- At the time of a user request

### Availability

- GEBCO Digital Atlas (CD):  
[http://www.bodc.ac.uk/projects/international/gebco/gebco\\_digital\\_atlas/](http://www.bodc.ac.uk/projects/international/gebco/gebco_digital_atlas/)
- GEBCO One Minute Grid (free): [http://www.bodc.ac.uk/data/online\\_delivery/gebco/](http://www.bodc.ac.uk/data/online_delivery/gebco/)
- User guide to the GEBCO One Minute Grid:  
[http://www.bodc.ac.uk/data/online\\_delivery/gebco/documents/gridhelp.pdf](http://www.bodc.ac.uk/data/online_delivery/gebco/documents/gridhelp.pdf)
- Software updates: [http://www.bodc.ac.uk/help\\_and\\_hints/software\\_updates/gebco.html](http://www.bodc.ac.uk/help_and_hints/software_updates/gebco.html)
- netCDF format documentation:  
[http://www.bodc.ac.uk/data/online\\_delivery/international\\_sea\\_level/woce\\_netcdf.html](http://www.bodc.ac.uk/data/online_delivery/international_sea_level/woce_netcdf.html)
- Cost: GEBCO Digital Atlas (CD): 230 pounds sterling<sup>481</sup>
- Terms and conditions: [http://www.bodc.ac.uk/data/online\\_delivery/gebco/copyright/](http://www.bodc.ac.uk/data/online_delivery/gebco/copyright/)
- Restrictions: Non-commercial reproduction in derivative form with accreditation<sup>482</sup>

### Comments

---

<sup>481</sup> <http://www.ngdc.noaa.gov/mgg/gebco/gebcocentenarycdrom.html>

<sup>482</sup> [http://www.bodc.ac.uk/data/online\\_delivery/gebco/copyright/](http://www.bodc.ac.uk/data/online_delivery/gebco/copyright/)

## Map Maker

Summary	
Type of tool	Mapping tool
Function	Resource display
Online / Desktop	Online
Computer infrastructure	Browser, JavaScript
Development status	Active
Time of use	At the time of a user request
Licence	



**Map Maker** is part of Australian National Resources Atlas and allows viewing a wide range of resource data displayed on a map on a selectable region of interest.

### Description

Map Maker has the ability to display many different mapping layers:

- Localities - 250K
- Aeronautical features
- Drainage - 250K
- State borders
- Roads - 250K
- Built up area - 250K
- Waterbodies - 250K
- Statistical Local Areas
- Natural Heritage Trust II boundaries
- Local Government Areas
- Australian Biogeographical Regionalisation
- Satellite image LANDSAT 100m
- Digital Elevation Model (DEM) 9 seconds
- Satellite image NOAA AVHRR 10km
- Tenure - 1996
- Land Use - 2001
- Coastline - 250K

Map Maker gives you the ability to customise a map from a second set of selectable layers

- Administrative / regional boundaries
- Agriculture
- Base layers
- Biodiversity and vegetation
  - Vegetation Cover
  - Protected Areas
  - National Vegetation Information System (NVIS) - Extant
  - National Vegetation Information System (NVIS) - Est. Pre-1750
- Climate
- Coasts
- Heritage
- Land
- Landscape Health

- People
- Rangelands
- Water

### Function

- Visualisation tools
  - Maps
- Non-biological data
  - Environmental data
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

- To display mapping layers of a range of resources

### Who will use this tool?

- Data users
  - Expert
  - Interest groups
- Limited skills required

### How will the tool be used?

- Online mapping tool
- Browser required with JavaScript
- User input required

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- At the time of a user request

### Availability

- Map Maker: <http://www.anra.gov.au/mapmaker/mapervlet?app=anra>
- ANRA: <http://www.anra.gov.au/>
- ANRA information products – Map Maker Layers: [http://www.anra.gov.au/topics/publications/fast\\_facts/fast-facts-33.html](http://www.anra.gov.au/topics/publications/fast_facts/fast-facts-33.html)
- Copyright: Commonwealth of Australia

### Comments

Map Maker is clunky to use:

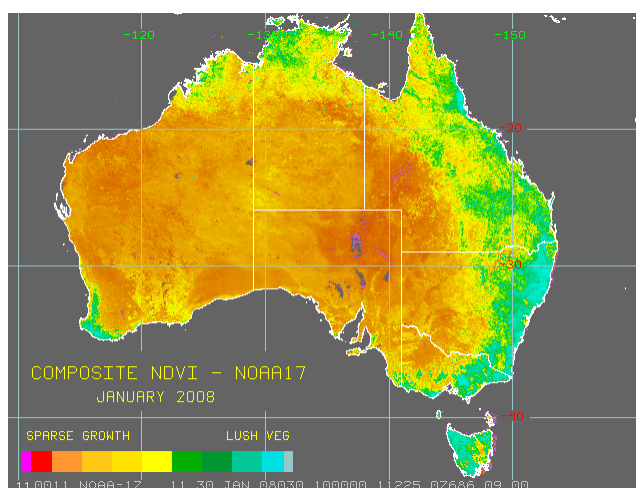
- It takes a long time to load (about 30+ seconds initially)
- It doesn't appear to cache (subsequent loads of the same map took just as long) and
- The program frequently crashes (server error).
- Text overwrites itself
- Text size is difficult to read
- Floating menu system is annoying (see Google Earth for another style of menu)

## NDVI – Normalised Difference Vegetation Index

Summary	
Type of tool	Archive of maps
Function	Vegetation maps
Online / Desktop	Online
Computer infrastructure	GIF images and GIS grids
Development status	Available monthly from 1997 to last month
Time of use	At the time of a user request
Licence	Access agreement

**Normalised Difference Vegetation Index (NDVI)** products are derived from Landsat images to allow the monitoring of the density and vigour of green vegetation growth.

### Description



NDVI image for January 2008.<sup>483</sup>

Normalised Difference Vegetation Index (NDVI) products are produced by the Bureau of Meteorology for the Australian region using measurements from the Advanced Very High Resolution Radiometer (AVHRR) on board the USA's NOAA polar orbiting meteorological satellites.<sup>484</sup>

This service provides access to images of Monthly Normalised Differential Vegetation Index Maps. **Monthly Composite Maps** are available from January 1997 to penultimate current month.<sup>485</sup>

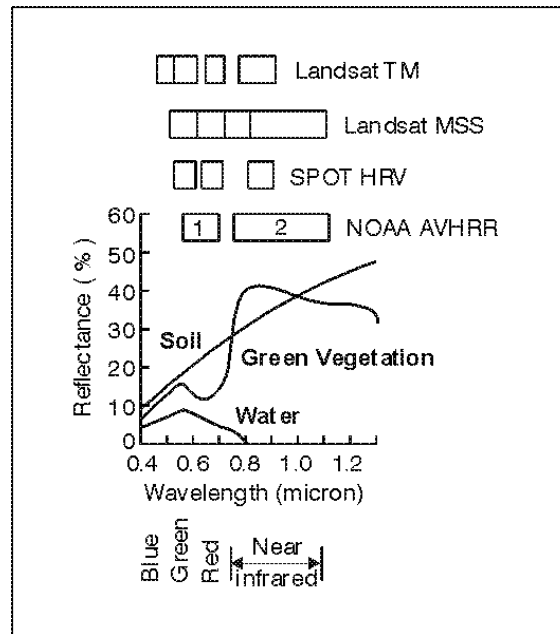
The differential reflectance in these bands provide a means of monitoring density and vigour of green vegetation growth using the spectral reflectivity of solar radiation. Green leaves commonly have larger reflectances in the near infrared than in the visible range. As the leaves come under water stress, become diseased or die back, they become more yellow and reflect significantly less in the near infrared range. Clouds, water, and snow have larger reflectances in the visible than in the near infrared while the difference is almost zero for rock and bare soil. Vegetation NDVI typically ranges from 0.1 up to 0.6, with higher values associated with greater

<sup>483</sup> <http://www.bom.gov.au/nmoc/NDVI/index.shtml>

<sup>484</sup> <http://www.bom.gov.au/nmoc/NDVI/index.shtml>

<sup>485</sup> <http://www.bom.gov.au/nmoc/NDVI/index.shtml>

density and greenness of the plant canopy. Surrounding soil and rock values are close to zero while the differential for water bodies such as rivers and dams have the opposite trend to vegetation and the index is negative. A range of errors such as scattering by dust and aerosols, Rayleigh scattering, subpixel-sized clouds, plus large solar zenith angles and large scan angles all act to increase Ch1 with respect to Ch2 and reduce the computed NDVI.



STORAGE/INT/CORF/UR3/14-1

*Spectral reflectance characteristics of common earth surface materials.*<sup>486</sup>

NOAA-14's radiometer is a 5-channel instrument which scans continuously at a maximum ground resolution of 1.1km and swath width of approximately 2,400 km. Local Area Coverage (LAC) data received in Melbourne is used to produce NDVI values from each orbit. Typically 2 sequential daytime orbits covering most of Australia are available for processing in near real-time each day. Each pixel is an average of the differential reflectance over the range of vegetation, soil types, water bodies and other surfaces within the pixel footprint.<sup>487</sup>

Other products include Monthly Maximum Value Composite (MVC) NDVI maps, Maximum Value Composite Differential (MVCD) NDVI products, Grassland Curing Index (GCI), Time series data, NDVI Grids in AXF format.

### Function

- Non-biological data
  - Environmental data
- User interface
  - Visual presentation

### Why use this tool?

- For historical changes in density and vigour of green vegetation

### Who will use this tool?

<sup>486</sup> Remote Sensing Digital Image Analysis, J.Richards,1986 from

<http://www.bom.gov.au/sat/NDVI/NDVI2.shtml>

<sup>487</sup> <http://www.bom.gov.au/sat/NDVI/NDVI2.shtml>

- Data users
  - Expert
  - Interest groups
- Special skills are required for GIS use

#### **How will the tool be used?**

- Online collection of images
- Maps: GIF format
- Grids for GIS: AXF McIDAS
- User input not required

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- At the time of a user request

#### **Availability**

- Bureau of Meteorology: <http://www.bom.gov.au/>
- NDVI products: <http://www.bom.gov.au/sat/NDVI/NDVI2.shtml>
- Regime of Access to BOM information: <http://www.bom.gov.au/other/regime1.shtml>
- Cost: generally freely available
- Licence: An access agreement may be necessary

#### **Comments**

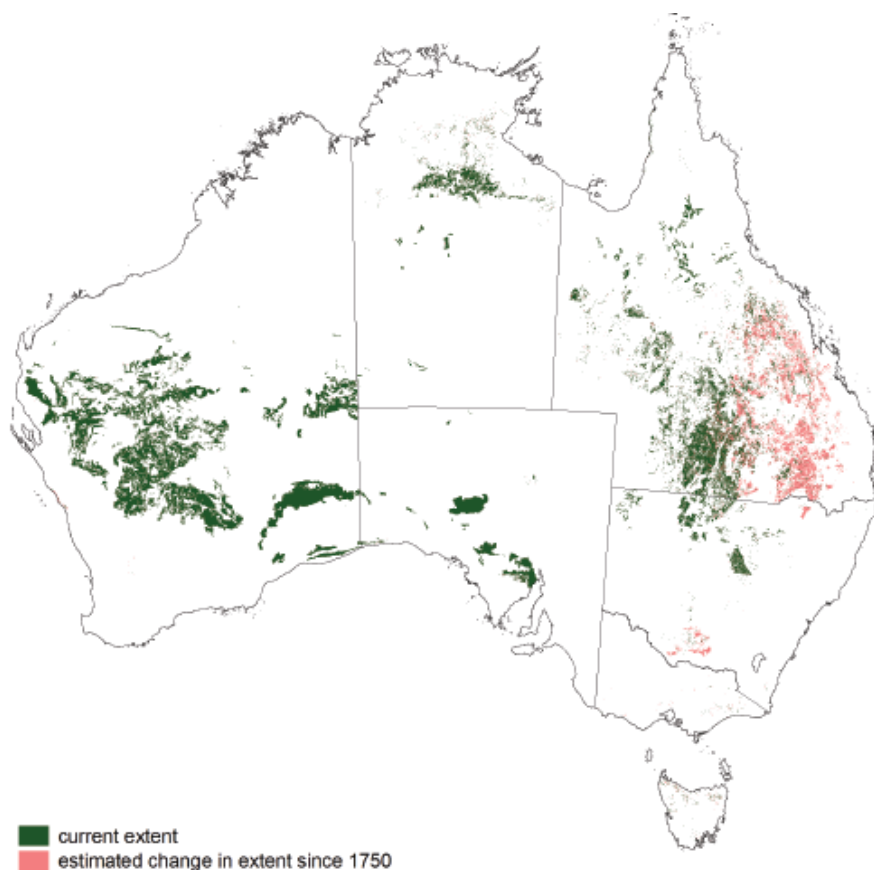
## NVIS – National Vegetation Information System

Summary	
Type of tool	Maps
Function	Viewing vegetation
Online / Desktop	Desktop
Computer infrastructure	GIS raster and vector datasets
Development status	NVIS Stage 1, Version 3.0 2005 Dataset, (2007)
Time of use	While data is with ALA. At time of user request
Licence	Non-commercial, personal or within organisation

**NVIS – National Vegetation Information System** is an Australian national database of vegetation data.

### Description

The NVIS framework was developed to enable the compilation of a nationally consistent vegetation dataset from data collected by states and territories. It provides a comprehensive means of describing and representing vegetation information based on establishing relationships between structural and floristic data. The NVIS Information Hierarchy is a system for describing the structural and floristic patterns of groups of plants in the landscape. Collectively, the different levels in the classification provide a description of vegetation that can be directly related to precise spatial areas as a vegetation map.<sup>488</sup>



*MVG6 - Acacia Forests and Woodland.*<sup>489</sup>

<sup>488</sup> <http://www.environment.gov.au/erin/nvis/about.html>

<sup>489</sup> DEW MVG Guide 2007



While many thousands of finer-level vegetation types have been drawn together into the NVIS database, they have been combined into broader groups to present a useful continental overview. The Australian Government Department of the Environment and Water Resources report has developed a set of 23 Major Vegetation Groups (MVGs) and 67 Major Vegetation Subgroups (MVSs) derived from the NVIS database, to enable a national view of native vegetation.<sup>490</sup>

Vegetation Theme	NVIS Stage 1, Version 3 (and fine-scaled gap-filling data)
Estimated pre-1750 vegetation	77%
Present vegetation	95%

*Coverage of detailed NVIS dataset.*<sup>491</sup>

A range of NVIS data products are available for use on different scales. Highly detailed data is received from state and territory custodians with standard NVIS attributes, and compiled into the national NVIS database. The resultant datasets are suited for use at a regional scale.<sup>492</sup>

The detailed data is also generalised and added to non-NVIS data to fill gaps, and recompiled to create the NVIS Major Vegetation Group (MVG) and Major Vegetation Subgroup (MVS) products. These generalised data are suited for national-scale analyses.<sup>493</sup>

There is also a 'whole-of-landscape' view of vegetation across Australia.<sup>494</sup>

### Function

- Visualisation tools
  - Maps
  - Images of representative Major Vegetation Groups
- Non-biological data
  - Environmental data
- User interface
  - Personal or institutional use
  - Visual presentation

### Why use this tool?

- A source of up-to-date and reliable vegetation information in Australia
- NVIS was developed to assist in managing a range of ecosystem services and practices such as biodiversity conservation, salinity control, improving water quality and fuel-load management.<sup>495</sup>

### Who will use this tool?

- Data providers
  - Institutions
- Data users

<sup>490</sup> DEW MVG Guide 2007

<sup>491</sup> DEW MVG Guide 2007

<sup>492</sup> <http://www.environment.gov.au/erin/nvis/data-products.html>

<sup>493</sup> <http://www.environment.gov.au/erin/nvis/data-products.html>

<sup>494</sup> <http://www.environment.gov.au/erin/nvis/data-products.html>

<sup>495</sup> <http://www.environment.gov.au/erin/nvis/about.html>

- Expert
- Interest groups
- ALA infrastructure

#### **How will the tool be used?**

- Raster and vector datasets are available for GIS – eg Raster dataset, 1 Ha cell size, ArcInfo Grid File
- Desktop use, data can be used locally
- Generally used in conjunction with other data

#### **Where in the data chain could this tool be used?**

- ALA central

#### **When could this tool be used?**

- While data is stored with ALA
- At the time of a user request

#### **Availability**

NVIS data products are available for use on different scales:

- National scale data: <http://www.environment.gov.au/erin/nvis/mvg/index.html>
- Regional scale data: <http://www.environment.gov.au/erin/nvis/detailed/index.html>
- Data is also available as a CD
- NVIS: <http://www.environment.gov.au/erin/nvis/>
- NVIS publications: <http://www.environment.gov.au/erin/nvis/publications>
- NVIS Stage 1, Version 3.0 2005 Dataset, (2007)
- Cost: free
- Copyright Department of Environment and Water Resources
- Licence: Non-commercial, personal use or within organisation  
see: [http://www.environment.gov.au/metadataexplorer/download\\_test\\_form.jsp](http://www.environment.gov.au/metadataexplorer/download_test_form.jsp)

#### **Comments**

## OzClim

Summary	
Type of tool	Climate prediction maps
Function	Modelling
Online / Desktop	Online
Computer infrastructure	Browser
Development status	New. Build 3.0.16 February 2008
Time of use	At the time of user request
Licence	Non-commercial use

Climate change scenarios for Australia.

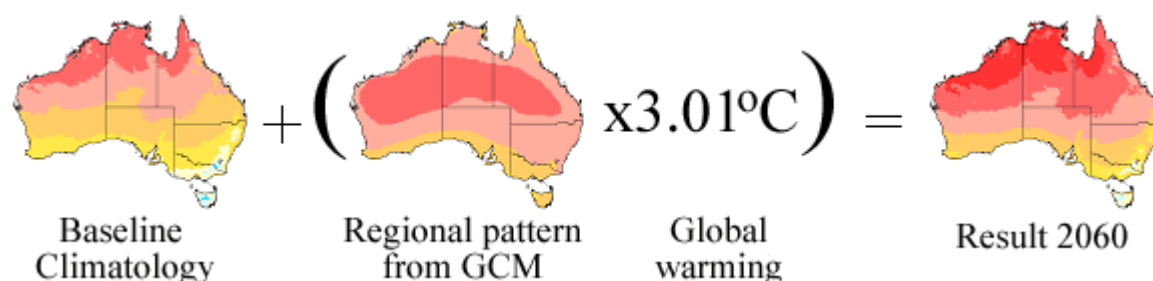
### Description

With **OzClim** you can:<sup>496</sup>

- generate climate change scenarios in a few easy steps
- explore temperature and rainfall climate scenarios from 2010 to 2100
- be guided through the process of generating your own climate scenarios
- download maps and scenario data for use in non-commercial research

OzClim provides a simple step-by-step option to help you generate and explore climate scenarios. There are also six scenarios in the examples section for rainfall and temperature for 2030.

The advanced section is designed for the scientific research community and policy making. Choose from eight climate models, eight emission scenarios and three climate sensitivities.<sup>497</sup>



*Climate scenario generation for annual temperature in 2060.*<sup>498</sup>

OzClim uses scenarios for greenhouse gases and sulfate aerosols from the IPCC Special Report on Emission Scenarios (SRES, 2000). Global warming projections derived from these scenarios in 5-yearly intervals were sourced from the IPCC (2001) Third Assessment Report.<sup>499</sup>

The SRES authors developed a series of storylines, based on assumptions about demographic change, economic development and technological advances, that were then given to modelling groups to estimate emissions for the major greenhouse gases and aerosols.<sup>500</sup>

<sup>496</sup> <http://www.csiro.au/ozclim/home.do>

<sup>497</sup> <http://www.csiro.au/ozclim/home.do>

<sup>498</sup> <http://www.csiro.au/ozclim/about.do>

<sup>499</sup> <http://www.csiro.au/ozclim/about.do#ProjectedGW>

<sup>500</sup> <http://www.csiro.au/ozclim/about.do#ProjectedGW>

OzClim is intended to provide:<sup>501</sup>

- increased accessibility to climate change scenarios for Australia, through a user-friendly web tool,
- a web based facility to help explain and generate scenarios for people new to climate change,
- advanced functionality for technical practitioners
- data on a 25 km grid over Australia, for 5-year increments in the 21st century, for user-selected greenhouse gas emission scenarios and climate models
- a facility to download the scenarios as text, images, or CSV files,
- data in a format suitable for input to impact studies.

Developed by CSIRO Australia.

### Function

- Non-biological data
  - Environmental data
- User interface
  - Personal and institutional use
  - Raw data and visual presentation

### Why use this tool?

The four major potential uses of OzClim are as:<sup>502</sup>

- a source of Australian climate scenario output
- a research tool to explore sectoral vulnerability to climate change for a wide range of initial assumptions
- a visualisation tool
- an educational tool

### Who will use this tool?

- Data users
  - Expert
  - Interest groups

### How will the tool be used?

- Online tool
- Export as image, excel file or GIS file
- User input required

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- At the time of a user request

### Availability

- OzClim: <http://www.csiro.au/ozclim/>
- CSIRO: <http://www.csiro.au>
- Science behind OzClim: <http://www.csiro.au/ozclim/about.do>

---

<sup>501</sup> <http://www.csiro.au/ozclim/about.do>

<sup>502</sup> <http://www.csiro.au/ozclim/about.do>

- Licence: Non-commercial use <http://www.csiro.au/ozclim/terms.do>

#### Comments

- These predicted climate maps could be used as inputs to other modelling tools attached to the ALA.
- A 25 km grid is too coarse for species modelling.<sup>503</sup>

---

<sup>503</sup> see: Chapman, A.D. (2004). Climate Surfaces for South America. Appendices E, F to /Sistema de Informação Distribuído para Coleções Biológicas: A Integração do Species Analyst e SinBiota. FAPESP/Biota process no. 2001/02175-5 March 2003 – March 2004./ Campinas, Brazil: CRIA  
[http://smlink.cria.org.br/docs/appendix\\_e.pdf](http://smlink.cria.org.br/docs/appendix_e.pdf), [http://smlink.cria.org.br/docs/appendix\\_f.pdf](http://smlink.cria.org.br/docs/appendix_f.pdf)

## WorldClim

Summary	
Type of tool	Application
Function	Mapping global climate
Online / Desktop	Desktop
Computer infrastructure	Used with GIS and in modelling tools such as Maxent
Development status	Version 1.4
Time of use	Post process
Licence	Educational and research only

**WorldClim** is a set of global climate layers (climate grids) with spatial resolutions from a square kilometre. They can be used for mapping and spatial modelling in a GIS or other computer program.<sup>504</sup>

### Description

WorldClim data layers cover the global land areas except Antarctica. They are in geodetic coordinate system (not projected, i.e., 'GEOGRAPHIC' or 'LATLONG' system). The datum is WGS84. They are available at 4 different spatial resolutions; from 30 seconds ( $0.93 \times 0.93 = 0.86 \text{ km}^2$  at the equator) to 2.5, 5 and 10 minutes ( $18.6 \times 18.6 = 344 \text{ km}^2$  at the equator).<sup>505</sup>

WorldClim was developed by Robert J. Hijmans, Susan Cameron, and Juan Parra, at the Museum of Vertebrate Zoology, University of California, Berkeley, in collaboration with Peter Jones and Andrew Jarvis (CIAT), and with Karen Richardson (Rainforest CRC).<sup>506</sup>

The data layers were generated through interpolation of average monthly climate data from weather stations on a 30 arc-second resolution grid (often referred to as "1 km<sup>2</sup>" resolution). Variables included are monthly total precipitation, and monthly mean, minimum and maximum temperature, and 19 derived bioclimatic variables.<sup>507</sup>

Climate change layers are also in preparation.<sup>508</sup>

### Function

- Analysis tools
- Visualisation tools
  - Maps

### Why use this tool?

- For spatially presenting climate data

### Who will use this tool?

- Data creation
- Data capture
- Data providers
  - Institutions

---

<sup>504</sup> <http://www.worldclim.org/>

<sup>505</sup> <http://www.worldclim.org/format.htm>

<sup>506</sup> <http://www.worldclim.org/>

<sup>507</sup> <http://www.worldclim.org/methods.htm>

<sup>508</sup> Arthur Chapman, Australian Biodiversity Information Services

- Private collections
  - Casual users
- Data users
  - Expert
  - General public

**How will the tool be used?**

- Species distribution modelling, climate change modelling<sup>509</sup>
- These datasets require a GIS or other program to display
- Can be used in conjunction with other mapping and modelling programs

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- As a post process, after data is with the user

**Availability**

- Download: <http://www.worldclim.org/>
- Licence: Free for academic and non-commercial use
- Version 1.4
- Future climate projections are expected

**Comments**

---

<sup>509</sup> Arthur Chapman, Australian Biodiversity Information Systems

## **Bibliography**



## AnimalBase

Summary	
Type of tool	Online library
Function	Historical literature reference
Online / Desktop	Online
Computer infrastructure	Web browser, pdf
Development status	Active
Time of use	At time of user request
Licence	Not for profit

**AnimalBase.** Early zoological literature online.

### Description

INSECTA COLEOPTERA. Scarabæus. 345

### I. COLEOPTERA.

*Elytra alas tegentia.*

170. SCARABÆUS. *Antennæ* clavatæ capitulo  
fiffili.  
*Tibiæ* anticæ sæpius dentatæ.

\* *Thorace* cornuto.

Hercules, 1. S. thoracis cornu incurvo maximo subtus barbato, ca-  
pitis cornu recurvato: supra dentato.

*Marcgr. bras.* 247. f. 3. *Jonst. inf.* t. 16. f. 1.  
*Olear. mus.* t. 16. f. 1. *Pet. gaz.* t. 70. f. 1.  
*Grew. mus.* 162. *Swamm. bibl.* t. 30. f. 2.  
*Raf. scarab.* 1. t. A. f. 1. *inf.* 4. p. 45. t. 5. f. 3.  
*Habitat in* America.

Aethon, 2. S. thorace bicorni, capitis cornu tridentato: apice bi-  
fido. *Mus. L. U.*

*Marcgr. bras.* 246. *Enena.* *Olear. mus.* t. 16. f. 2.

*Linnæus, C. 1758. Systema naturæ per regna tria naturæ, secundum classes, ordines, genera, species, cum characteribus, differentiis, synonymis, locis. Tomus I. Editio decima, reformata. - pp. [1-4], 1-824. Holmiæ. (Laurentii Salvii).<sup>510</sup>*

Our objective is to provide free access for all scientists to the old zoological literature, particularly to those important publications where name-bearing zoological taxa were originally described. The literature is digitized in image format by the [SUB Göttingen](#) (our university library). In a first 2-year period (2003-2005) financed by the DFG we have digitized nearly all taxonomically relevant zoological literature from the beginnings until 1770 (about 400 works). Only some 5 % of the literature is not present in Göttingen and we are currently trying to obtain some works from other libraries.

In a second 2-year-period we will try to cover the period from 1770 to 1800. Monographic works and journal articles shall both be digitized.

<sup>510</sup> [http://gdz.sub.uni-goettingen.de/no\\_cache/dms/load/img/](http://gdz.sub.uni-goettingen.de/no_cache/dms/load/img/)

The AnimalBase database is primarily established to link the old literature with the names of the animals described therein. We have continuously gone through the old works, from 1757 onwards, and entered all correctly described new animal names (genera and species taxa) by hand according to a standard established by our working group.

Today our database should include all zoological taxa described from 1757 until 1770. We are continuously progressing in the time-line. The names are entered into the database along the original literature and largely independently from other databases and secondary literature sources, so that errors will not be duplicated.

This is a public and non-profit service.<sup>511</sup>

### **Function**

- Non-electronic data
  - literature
  - digitisation
- User interface
  - Personal use
  - Visual presentation

### **Why use this tool?**

Users from all over the world have free access to the complete database information and to all digitized pages of the old zoological literature.<sup>512</sup>

### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators
- Data users
  - Expert
- No special skills required

### **How will the tool be used?**

- Online bibliographic database search

Search options:

- species and genus taxa according to original description
- species, genera according to current classification
- families
- higher groups

### **Where in the data chain could this tool be used?**

- User's machine

### **When could this tool be used?**

- At the time of a user request

---

<sup>511</sup> <http://www.animalbase.de/>

<sup>512</sup> <http://www.animalbase.de/>

### Availability

- AnimalBase: <http://www.animalbase.de/>
- AnimalBase standard: <http://www.animalbase.uni-goettingen.de/zooweb/animalbase-standard-col.html;jsessionid=BA370A981E587B6C2E21409C83FEE829>
- Active November 2007
- Restrictions: Not for profit<sup>513</sup>

### Comments

---

<sup>513</sup> <http://www.animalbase.de/>

## Biodiversity Heritage Library

Summary	
Type of tool	Online library
Function	Historical literature reference
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Active
Time of use	At time of a user request
Licence	Open access as part of global Biodiversity Commons

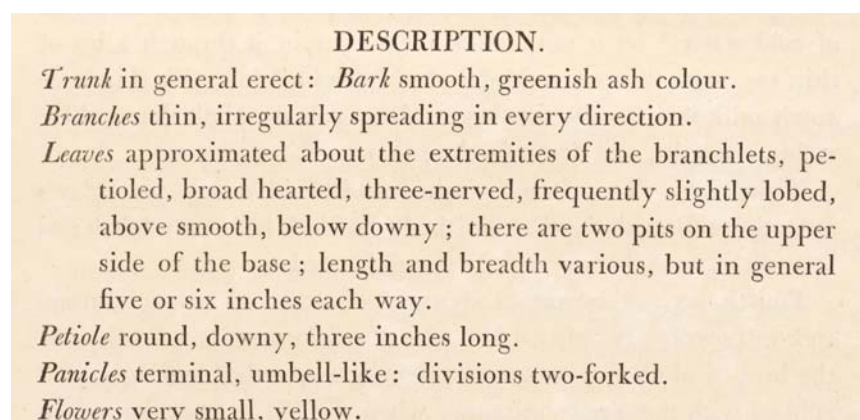
The **Biodiversity Heritage Library** is a digitisation project giving web searchable access to a large collection of published biodiversity literature.

### Description

The Natural History Museum in London and nine other American major natural history museum libraries, botanical libraries, and research institutions have joined to form the Biodiversity Heritage Library. BHL partners will digitize the published literature of biodiversity held in their respective collections, providing basic, important content for immediate research and for multiple bioinformatics initiatives. For the first time in history, the core of our natural history and herbaria library collections will be available to a truly global audience.<sup>514</sup>

The participating libraries have over two million volumes of biodiversity literature collected over 200 years.<sup>515</sup> About 1.4 million pages were online in November 2007.

A characteristic of the Biodiversity Heritage Library that distinguishes it from other mass digitization projects is the incorporation of service-based algorithms to identify scientific name strings throughout digitized content. These 'taxonomically intelligent' services, powered by uBio.org's TaxonFinder and NameBank, have been incorporated into the BHL Portal to provide names-based interfaces into taxonomic literature.<sup>516</sup>



Extract from *Plants of the coast of Coromandel*, by William Roxburgh and Joseph Banks.<sup>517</sup>

<sup>514</sup> <http://biodiversitylibrary.blogspot.com/2007/11/discovered-bibliographies-through.html>

<sup>515</sup> <http://www.biodiversitylibrary.org/About.aspx>

<sup>516</sup> <http://biodiversitylibrary.blogspot.com/2007/11/discovered-bibliographies-through.html>

<sup>517</sup> <http://www.biodiversitylibrary.org/item/9711>

### How it works

Each digitized page image in BHL has an accompanying OCR text file. As users navigate to a page, the uncorrected OCR file is sent to uBio's TaxonFinder, which identifies text strings that match the characteristics of Latin binomials. Those potential name strings are then compared to the 10.7 million+ names in uBio's NameBank, and the results, both matched and unmatched, are stored in the BHL database. BHL also has automated processes to reindex pages at regular intervals since NameBank is a growing repository.<sup>518</sup>

### Function

- Bibliographic reference
- Taxonomy
- Non-electronic data
  - literature
  - digitisation
- User interface
  - Personal and institutional use

### Why use this tool?

- To sight a direct copy and use a text version of the original document.
- Keyword search of historical biodiversity literature

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators – specimens, identification
- Data users
  - Expert
  - Interest groups
  - General public

### How will the tool be used?

- Web based online tool, user input required
- Taxonomic Name Service are XML-based web services that can be invoked via SOAP or HTTP GET/POST requests.<sup>519</sup>

### Where in the data chain could this tool be used?

- Data source
- User's machine

### When could this tool be used?

- Before data is made available to ALA
- At the time of a user request
- As a post process, after data is with the user

### Availability

- BHL: <http://www.biodiversitylibrary.org/>
- RSS feed: <http://biodiversitylibrary.blogspot.com/feeds/posts/default?alt=rss>

---

<sup>518</sup> <http://biodiversitylibrary.blogspot.com/2007/11/discovered-bibliographies-through.html>

<sup>519</sup> <http://biodiversitylibrary.blogspot.com/2007/12/name-services-last-updated-12052007-cdf.html>

- Name search: <http://www.biodiversitylibrary.org/NameSearch.aspx>
- Creative Commons Licence: <http://www.biodiversitylibrary.org/Copyright.aspx>
- Free for non-commercial use. Commercial licensing available: [bhl@si.edu](mailto:bhl@si.edu)
- Example of how to use the BHL with a web link: Banksia serrata, Old Man Banksia  
[http://www.biodiversitylibrary.org/name/Banksia\\_serrata](http://www.biodiversitylibrary.org/name/Banksia_serrata)

### Comments

These results are generated from uncorrected OCR, which range in quality from pretty good (contemporary publications, such as modern issues of *Rhodora*) to downright terrible (18th century Latin texts, such as *Species Plantarum*). Again, further evaluation is required to determine the full scope of this problem.<sup>520</sup>

### In development

Up next is development of a service layer on top of the names index so that other application providers can query & display 'discovered bibliographies' within their own applications. This service will be deployed in early 2008.<sup>521</sup>

This is a largely American/British project, and may not have sufficient Australian species references.

Is this an opportunity for a local effort to digitise some of the published biodiversity literature of interest locally?

---

<sup>520</sup> <http://biodiversitylibrary.blogspot.com/2007/11/discovered-bibliographies-through.html>

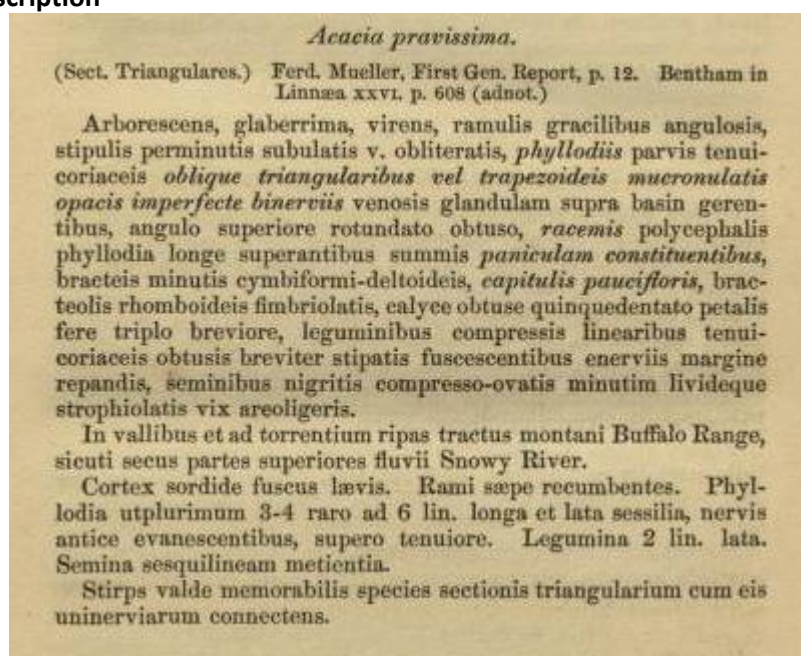
<sup>521</sup> <http://biodiversitylibrary.blogspot.com/2007/11/discovered-bibliographies-through.html>

## Botanicus

Summary	
Type of tool	Online historical literature
Function	Botanical taxonomy and bibliography
Online / Desktop	Online
Computer infrastructure	Browser, pdf, searchable text
Development status	Advanced. 1700 volumes, 800,000 pages digitised
Time of use	Before data is available for the ALA. At the time of a user request.
Licence	Free for non-commercial use

**Botanicus** is a freely accessible, web-based encyclopedia of historic botanical literature from the Missouri Botanical Garden Library.<sup>522</sup>

### Description



*Fragmenta phytographiæ Australiæ* Ferdinandus Mueller, 1858-82.<sup>523</sup>

Over 67,000 systematic botanical publications exist, but only those most recently published are in digitized form.<sup>524</sup> Botanicus takes the 18th and 19th century botanical literature and captured it as images along with associated metadata. OCR is used, and XML mark-up protocols create text and embedded links for external use. Name descriptions are found through TROPICOS, and botanical names are indexed through uBio.

The [Missouri Botanical Garden Library](#) has been digitizing materials since 1995, focusing primarily on beautifully illustrated volumes from their rare book collection. Botanicus is funded by the W.M. Keck Foundation, and expands the selection criteria to include non-illustrated works of significant importance to taxonomic botany.<sup>525</sup> The literature to be digitized in the

<sup>522</sup> <http://www.botanicus.org/browse>

<sup>523</sup> <http://www.botanicus.org/title/b11906789>

<sup>524</sup> <http://www.botanicus.org/About.aspx>

<sup>525</sup> <http://www.botanicus.org/About.aspx>

course of this project will be selected and prioritized based on the frequency of use and scientific importance of the bibliographic references.<sup>526</sup>

### Function

- Taxonomy
- Non-electronic data
  - digitisation
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

Digitizing, indexing, and annotating historical scientific literature is vital to future research in systematic botany. Like other natural history disciplines - but unlike the physical sciences - systematic botany is built upon and requires frequent reference to the literature of its past. To conduct carefully documented and authenticated research, botanists must spend weeks in library collections searching the published botanical literature for data to develop a new project or substantiate their recent observations. Comprehensive collections of botanical literature are only available in a handful of libraries, all located in North American and Europe.

For botanical researchers, these library-centered literature searches, while a crucial requirement of any project, delay hypothesis development or recognition and publication of new plant discoveries. For those traveling in remote parts of North America or stationed overseas, lack of access to library resources compounds these difficulties. Further, no matter how scrupulous the search, when scientists must work manually through an array of journals and books it is impossible to be sure that all historical facts have been located and all published observations have been seen.<sup>527</sup>

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators
- Data users
  - Expert
- Foreign language skills may be required

### How will the tool be used?

- Online
- Scanned images are in pdf
- User input required

Botanicus uses *taxonomic intelligence* tools, including TaxonFinder developed by uBio.org, to locate, verify, and record scientific names located within the text of each digitized page. You can return a list of every page on which Botanicus has encountered a particular scientific name.

**Note:** The text used for this identification is uncorrected OCR, so may not include all results expected or visible in the page.<sup>528</sup>

---

<sup>526</sup> <http://www.botanicus.org/ViewTitles.aspx>

<sup>527</sup> <http://www.botanicus.org/About.aspx>

<sup>528</sup> <http://www.botanicus.org/NameSearch.aspx>



**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request

**Availability**

- Botanicus: <http://www.botanicus.org/>
- Missouri Botanical Garden Library: <http://www.mobot.org/mobot/molib/>
- Taxonomic name search: <http://www.botanicus.org/NameSearch.aspx>
- Recent additions RSS feed: <http://www.botanicus.org/RecentRss.aspx>
- Cost: free
- 1700 volumes, 800,000 pages digitised as of February 2008
- Copyright: free for non-commercial use
- Licence: [Creative Commons Attribution-Noncommercial 2.5](#)

**Comments**

This has a very similar look and feel to the [Biodiversity Heritage Library](#). The Missouri Botanical Garden Library is a contributor to the BHL.

Like the BHL, Botanicus has limited numbers of books of direct interest to Australian Botany.

## BUGZ – Bibliography of New Zealand Terrestrial Invertebrates

Summary	
Type of tool	Online database
Function	Bibliographic reference
Online / Desktop	Online
Computer infrastructure	Browser, pdf
Development status	Active
Time of use	At time of user request
Licence	Personal use, not for redistribution

**BUGZ** is a full-text electronic archive of the *Bibliography of New Zealand Terrestrial Invertebrates 1775 – 1993*.<sup>529</sup>

### Description

BUGZ is a user-friendly web interface designed to allow full-text search and retrieval of information from New Zealand's largest compilation of invertebrate literature – the 'BUGS' bibliography (Ramsay & Crosby 1992). 'BUGZ' contains a literature database of 16,080 articles on the terrestrial invertebrates of New Zealand, published between 1775 and 1993 and provides full-text indexing of the more than 200,000 pages of text scanned from the articles of the BUGS bibliography. This massively enhances the search capabilities and subsequent access to archived information on the taxonomic status, life history, ecology, and conservation significance in the primary literature on New Zealand's terrestrial invertebrates.

Apart from the ability to undertake full-text searching, BUGZ is the first New Zealand biodiversity database to allow dynamic matching of its entire full-text database against the taxonomic namebank of uBio – the universal Biological indexer and organiser.

All written sources in the public domain were included in the BUGS bibliography, including both core scientific papers and popular accounts. Ramsay and Crosby (1992) estimated that about 95% of the core literature and at least 60% of the popular literature was included for the major invertebrate groups.

BUGZ was a collaborative project involving the University of Canterbury (School of Biological Sciences), Landcare Research and uBio.<sup>530</sup>

### Function

- Non-electronic data
  - literature
  - digitisation
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

---

<sup>529</sup> <http://entdocs.landcareresearch.co.nz/WebForms/about.aspx>

<sup>530</sup> <http://entdocs.landcareresearch.co.nz/WebForms/people.aspx>

Online access to well indexed scientific publications is an extremely efficient use of those references and removes much of the difficulty in retrieving information from older publications.

- See also Biodiversity Heritage Library <http://www.biodiversitylibrary.org/>

#### Who will use this tool?

- Data users
  - Expert
  - Interest groups
  - General public
- No special skills required

#### How will the tool be used?

- Online database
- Full text searching, not just for species names
- Matching of the text against the uBio taxonomic namebank
- Scanned images of original articles (copyright permitting)
- Downloadable pdf of high quality scanned image of original articles (copyright permitting)
- User input is required

#### Where in the data chain could this tool be used?

- User's machine

#### When could this tool be used?

- At the time of a user request

#### Availability

- BUGZ: <http://entdocs.landcareresearch.co.nz/WebForms/SearchForm.aspx?doSearch=false>
- Help: <http://entdocs.landcareresearch.co.nz/WebForms/help.aspx>
- Contact: [Raphael.Didham@canterbury.ac.nz](mailto:Raphael.Didham@canterbury.ac.nz)
- Cost: free
- Copyright is owned by the original publisher
- Restrictions: no reproduction except for personal download and printing for personal use<sup>531</sup>

#### Comments

BUGZ has tackled the difficult problem of copyright by allowing access to articles in a variety of ways. Depending on the copyright restrictions, an article may be available as:

1. A reference only
  2. A scanned version of the article presented as a set of embedded jpg pages
  3. A pdf file of a high quality scanned image of the article which can be downloaded for private use only and not for publication
- The full-text version is not accessible for downloading.<sup>532</sup>
  - The scanning quality is often exceptional, given the age of many of the documents.
  - Although this database is explicitly New Zealand focussed, it will have much use in an Australian context, for example for those interested in insect pest species.

---

<sup>531</sup> <http://entdocs.landcareresearch.co.nz/WebForms/Disclaimer.aspx>

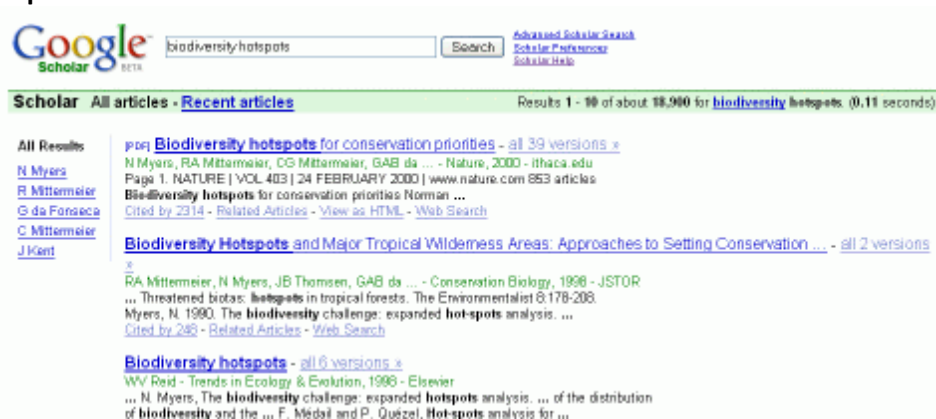
<sup>532</sup> <http://entdocs.landcareresearch.co.nz/WebForms/about.aspx>

## Google Scholar

Summary	
Type of tool	Application
Function	Bibliographic service
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Beta version
Time of use	At time of user request
Licence	

Google Scholar provides a search of scholarly literature across many disciplines and sources, including theses, books, abstracts and articles.<sup>533</sup>

### Description



Google Scholar provides a simple way to broadly search for scholarly literature. From one place, you can search across many disciplines and sources: peer-reviewed papers, theses, books, abstracts and articles, from academic publishers, professional societies, preprint repositories, universities and other scholarly organizations. Google Scholar helps you identify the most relevant research across the world of scholarly research.

### Rank

Google Scholar aims to sort articles the way researchers do, weighing the full text of each article, the author, the publication in which the article appears, and how often the piece has been cited in other scholarly literature. The most relevant results will always appear on the first page.<sup>534</sup>

### Function

- Bibliography
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

- What is the purpose of the tool?
- Are there alternatives?

<sup>533</sup> <http://scholar.google.com.au/>

<sup>534</sup> <http://scholar.google.com.au/intl/en/scholar/about.html>

### Who will use this tool?

- Data creation
  - Experts
- Data capture
  - Curators
- Data users
  - Expert
  - Interest groups

### How will the tool be used?

- Online bibliographic search tool
- User input required
- Simple and advanced search

Advanced search allows for searching by:<sup>535</sup>

- word order and location
- author
- publication
- date
- subject areas



Google Scholar can also be implemented on your website, allowing users to find scholarly papers hosted locally.<sup>536</sup>

### Where in the data chain could this tool be used?

- User's machine

### When could this tool be used?

- At the time of a user request

### Availability

- Google Scholar: <http://scholar.google.com.au/>
- Advanced Scholar Search: [http://scholar.google.com.au/advanced\\_scholar\\_search](http://scholar.google.com.au/advanced_scholar_search)
- Help: <http://scholar.google.com.au/intl/en/scholar/help.html>
- Cost: free

### Comments

- Google Scholar links to only a subset of available literature.
- A search result will often take you to a publisher's website, where the full text article can be purchased.

---

<sup>535</sup> [http://scholar.google.com.au/advanced\\_scholar\\_search](http://scholar.google.com.au/advanced_scholar_search)

<sup>536</sup> [https://services.google.com/inquiry/scholar\\_email](https://services.google.com/inquiry/scholar_email)

## PubMed

Summary	
Type of tool	Database
Function	Bibliographic literature
Online / Desktop	Online
Computer infrastructure	Browser, RSS available, some batch ability
Development status	Current
Time of use	At the time of a user request
Licence	Public domain. Free for re-distribution

Through **PubMed**, you can search 16,000,000 biomedical journal abstracts. PubMed Central is a database that contains whole research articles from over 300 research journals.<sup>537</sup>

### Description

**Entrez** is the text-based search and retrieval system used at NCBI for services including **PubMed**, Nucleotide and Protein Sequences, Protein Structures, Complete Genomes, Taxonomy, OMIM, and many others. PubMed provides access to citations from biomedical literature. LinkOut provides access to full-text articles at journal Web sites and other related Web resources. PubMed also provides access and links to the other Entrez molecular biology resources.<sup>538</sup>

Publishers participating in PubMed electronically submit their citations to NCBI prior to or at the time of publication. If the publisher has a web site that offers full-text of its journals, PubMed provides links to that site as well as biological resources, consumer health information, research tools, and more. There may be a charge to access the text or information.

PubMed, available via the NCBI Entrez retrieval system, was developed by the National Center for Biotechnology Information (NCBI) at the National Library of Medicine (NLM), located at the U.S. National Institutes of Health (NIH).<sup>539</sup>

### Function

- Taxonomy
- Bibliographic
- User interface
  - Personal and institutional use
  - Raw data
  - abstracts and some full text

### Why use this tool?

- Bibliographic references

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Data capture
  - Curators

---

<sup>537</sup> <http://www.ncbi.nlm.nih.gov/Literature/>

<sup>538</sup> <http://www.ncbi.nlm.nih.gov/entrez/query/static/overview.html>

<sup>539</sup> <http://www.ncbi.nlm.nih.gov/entrez/query/static/overview.html>

- Data providers
  - Institutions
  - Private collections
- Data users
  - Expert
  - Interest groups
- Online library skills required

#### How will the tool be used?

- Online bibliographic database
- The Batch Citation Matcher will match citations to PubMed using bibliographic information such as journal, volume, issue, page number, and year, or the Entrez Programming Utilities can provide access to Entrez data outside of the regular Web query interface.<sup>540</sup>
- Linking PubMed to a website: <http://www.ncbi.nlm.nih.gov/entrez/query/static/link.html>
- Entrez Programming Utilities are tools that provide access to Entrez data outside of the regular web query interface:
  - [http://eutils.ncbi.nlm.nih.gov/entrez/query/static/eutils\\_help.html](http://eutils.ncbi.nlm.nih.gov/entrez/query/static/eutils_help.html)
- Creating a Web Link to the Entrez Databases: (see example in **Comments** below)
  - <http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=helplinks.chapter.linkshelp>
- PubMed on handheld devices:
  - <http://www.ncbi.nlm.nih.gov/entrez/query/static/faq.html#handheld>
- Citations can be exported into a reference manager program:
  - [http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=helppubmed.section.pubmedhelp.Saving\\_and\\_Emailing\\_#pubmedhelp.Exporting\\_citations](http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=helppubmed.section.pubmedhelp.Saving_and_Emailing_#pubmedhelp.Exporting_citations)
- How to get a full text article:
  - [http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=helppubmed.section.pubmedhelp.How\\_to\\_Get\\_the\\_Journ](http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=helppubmed.section.pubmedhelp.How_to_Get_the_Journ)
- User input is required
- Some batch ability

#### Where in the data chain could this tool be used?

- User's machine

#### When could this tool be used?

- At the time of a user request

#### Availability

- PubMed: <http://www.pubmed.gov/>
- PubMed Central: <http://www.pubmedcentral.nih.gov/>
- FAQ: <http://www.ncbi.nlm.nih.gov/entrez/query/static/faq.html>
- PubMed Help:
  - <http://www.ncbi.nlm.nih.gov/books/bv.fcgi?rid=helppubmed.chapter.pubmedhelp>
- MEDLINE factsheet: <http://www.nlm.nih.gov/pubs/factsheets/medline.html>
- Cost: free<sup>541</sup>
- Restrictions: US Government information available from PubMed site is within the public domain. Public domain information on the National Library of Medicine (NLM) Web pages

<sup>540</sup> <http://www.ncbi.nlm.nih.gov/entrez/query/static/overview.html>

<sup>541</sup> <http://www.nlm.nih.gov/databases/leased.html>

may be freely distributed and copied. However, it is requested that in any subsequent use of this work, NLM be given appropriate acknowledgment.

### **Comments**

What's the difference between MEDLINE and PubMed?

[http://www.nlm.nih.gov/pubs/factsheets/dif\\_med\\_pub.html](http://www.nlm.nih.gov/pubs/factsheets/dif_med_pub.html)

Example of a web link to return the first 50 citations for *Thylacine* in 2006:

<http://www.ncbi.nlm.nih.gov/sites/entrez?cmd=search&db=pubmed&term=thylacine+AND+2006&dispmax=50>



## Field Data Collection

## ArcPad

Summary	
Type of tool	Mobile GIS software
Function	Field data collection
Online / Desktop	Mobile desktop
Computer infrastructure	ESRI GIS
Development status	Commercial
Time of use	Before data is made available to the ALA. At time of user request.
Licence	Single use

**ArcPad** is software for mobile GIS and field mapping applications using handheld and mobile devices. ArcPad provides field-based personnel with the ability to capture, analyse, and display geographic information.<sup>542</sup>

### Description



**ArcPad** allows users to create and edit spatial data (points, lines, and polygons) using input from the mouse pointer, pen, GPS, rangefinder, or camera. ArcPad includes editing tools for offsets, repeated attributes, segmented line features, snapping, and more.

Users can record field sketches and notes into the data collection process with ArcPad. Sketches can be shared and exchanged as a graphics layer to and from ArcGIS Desktop. These sketches can become part of the geodatabase and the project record.

A key feature of ArcPad is the ability to display a variety of geographic and attribute information. ArcPad uses data directly from an individual's desktop or an organization's enterprise GIS without the need to convert to unique portable formats. ArcPad supports vector and raster data in a multilayered environment. Users can combine vector and raster data with the only limitations being the speed and memory capacity of the hardware in use.<sup>543</sup>

<sup>542</sup> <http://www.esri.com/software/arcgis/arcpad/>

<sup>543</sup> <http://www.esri.com/software/arcgis/arcpad/about/data-collection.html>

ArcPad supports field devices, such as GPS receivers, rangefinders, and integrated digital cameras with GIS data collection. Input from these devices is stored within the GIS.<sup>544</sup>

### Function

- Visualisation tools
  - Maps
  - Images
- Provider interaction
  - Data preparation
- Non-biological data
  - Site specific data
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

- As a replacement for paper-based field notebooks
- To bring a limited computer desktop GIS functionality to the field.

### Who will use this tool?

ArcPad is designed for organizations wanting to expand the benefits of GIS to workers in the field.<sup>545</sup>

- Data creation
- Data capture
- Data providers
  - Casual users
- GIS skills required

### How will the tool be used?

ArcPad supports the following data formats:<sup>546</sup>

- ArcPad AXF file (based on Microsoft SQL Server Compact Edition)
- ESRI shapefile
- ArcPad graphics layers
- ArcPad photo layers
- MrSID MG2 and MG3
- JPEG
- JPEG2000
- TIFF
- GIF
- PNG
- Windows bitmap
- CADRG
- ArcIMS Image Services via access to the Internet
  
- Stand-alone tool for use on mobile devices
- User input required

---

<sup>544</sup> <http://www.esri.com/software/arcgis/arcpad/about/features.html>

<sup>545</sup> <http://www.esri.com/library/brochures/pdfs/arcpadbro.pdf>

<sup>546</sup> <http://www.esri.com/software/arcgis/arcpad/about/data-collection.html>

Attribute data is descriptive information about a location and can be edited and collected with ArcPad data collection forms. These forms include fields that are designed for input using a variety of devices (Windows Mobile devices, laptops, Tablet PCs). Forms include field options for text boxes, menus, calendar/date selection, and more. Simple data collection forms can be created quickly with ArcPad from a shapefile. More specialized custom forms can be created with ArcPad Application Builder and deployed on ArcPad.<sup>547</sup>

**Where in the data chain could this tool be used?**

- Data source

**When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request

**Availability**

- ArcPad: <http://www.esri.com/software/arcgis/arcpad/>
- ESRI: <http://www.esri.com/>
- FAQ: <http://www.esri.com/software/arcgis/arcpad/about/faqs.html>
- Cost: US\$495. For local pricing: <http://www.esri.com/software/arcgis/arcpad/how-to-buy.html>
- Single use licence

**Comments**

---

<sup>547</sup> <http://www.esri.com/software/arcgis/arcpad/about/data-collection.html>

## EFG – Electronic Field Guide

Summary	
Type of tool	Identification key
Function	Taxonomy
Online / Desktop	Online
Computer infrastructure	Browser, Java
Development status	Active
Time of use	As a pre-process, or as a post-process
Licence	Free Software Foundation General Public Licence

**EFGs – Electronic Field Guides** use technology to enhance the species identification process. Each identification key can be accessed in a variety of ways: navigate through a Java-based picture key, a printable text-only key, or "Browse All" to see all the images on a single page.<sup>548</sup>

### Description

The EFG project aims to provide a database management system for manipulating taxonomic information associated with the tree of life based on Linnaeus' model. It is intended for use as an educational research tool to assist users in accessing data regarding taxa. It will help biology researchers to identify species by performing iterative classification through elaboration of characteristics.<sup>549</sup>

EFG helps biology researchers to identify species by performing iterative classification through elaboration of characteristics.

A browse feature provides an interface for the navigation of the Tree of Life. You could browse in the following two ways:<sup>550</sup>

1. By Direct Taxon: To request information on a specific taxon by typing in its name.
2. By Branch Traversal: To traverse the tree of life using a graphical interface.

A search feature provides an interface for searching the tree of life on the basis of the search criteria submitted by the user. The search criteria can be based on properties such as colour and shape of wings.<sup>551</sup>

From a single spreadsheet file and a set of images, the efgKEYS software can render the key in a variety of formats, from web-based to printable pdf to a version you can upload to your iPod.<sup>552</sup>

This project is the result of a collaborative effort between the Departments of Computer Science and Biology at the University of Massachusetts Boston, with funding from the National Science Foundation.<sup>553</sup>

### Function

- Visualisation tools
  - Images

---

<sup>548</sup> <http://efg.cs.umb.edu/keys/html/html/welcome.html>

<sup>549</sup> <http://efg.cs.umb.edu/keys/html/html/welcome.html>

<sup>550</sup> <http://efg.cs.umb.edu/keys/html/html/overview.html>

<sup>551</sup> <http://efg.cs.umb.edu/keys/html/html/overview.html>

<sup>552</sup> <http://efg.cs.umb.edu/efg/efgsoftware.html>

<sup>553</sup> <http://efg.cs.umb.edu/efg/>

- Taxonomy
  - Identification tools, keys
- Provider interaction
  - Data preparation
- User interface
  - Personal use
  - Visual presentation

#### **Why use this tool?**

- To build or use an interactive, electronic, tree-based and multiple entry key

#### **Who will use this tool?**

- Data creation
  - Experts - taxonomy
- Data users
  - Expert
  - Interest groups
  - General public
- Special taxonomic skills are required to create keys

#### **How will the tool be used?**

- EFG keys are created using an Excel-based interface<sup>554</sup>
- Example of Excel file: <http://efg.cs.umb.edu/keys/html/html/Ithomids.xls>
- Key access online - software is available to download and serve these keys yourself
- Java required
- User input required

#### **Where in the data chain could this tool be used?**

- Data source
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- As a post process, after data is with the user

#### **Availability**

- EFG – Electronic Field Guide: <http://efg.cs.umb.edu/keys/html/index.html>
- How to create a key: <http://efg.cs.umb.edu/keys/html/html/KeyAuthorInstructions.html>
- Downloads: <http://efg.cs.umb.edu/EFGsoftware/index.html>
- Contact: robert.stevenson @umb.edu
- Cost: free
- Licence: The EFG software may be used and copied under the terms of the Free Software Foundation [General Public License](#).<sup>555</sup>

#### **Comments**

---

<sup>554</sup> <http://bdei.cs.umb.edu/>

<sup>555</sup> <http://efg.cs.umb.edu/keys/html/html/welcome.html>

## TDS Nomad

Summary	
Type of tool	Rugged handheld computer
Function	Field data collection
Online / Desktop	Stand-alone
Computer infrastructure	Windows mobile
Development status	Released. Model 800 (2007)
Time of use	At time of collection. Before data is made available to ALA
Licence	

**TDS Nomad** is a rugged handheld computer built for field data collection.<sup>556</sup>

### Description



557

The Nomad is made to collect, save and transmit data in the field.

This is a handheld computer with Windows Mobile software, able to perform standard MS Office functions and data collection.

Features:<sup>558</sup>

- Inbuilt GPS
- 1 GB flash storage
- Colour VGA display
- Optional colour digital camera, 2 megapixel
- Optional 1D laser barcode scanner
- CF and SDIO slots
- MIL-STD-810F military specification for drops and ruggedness
- IP67 dust and water sealed
- USB connection

### Function

- Data collection
- Visualisation tools
  - Maps
- Provider interaction
  - Data preparation
- User interface

<sup>556</sup> <http://www.tdsway.com/products/nomad/overview>

<sup>557</sup> [http://www.tdsway.com/binaries/nomad/docs/TDS\\_Nomad\\_data\\_sheet.pdf](http://www.tdsway.com/binaries/nomad/docs/TDS_Nomad_data_sheet.pdf)

<sup>558</sup> See [http://www.tdsway.com/binaries/nomad/docs/TDS\\_Nomad\\_data\\_sheet.pdf](http://www.tdsway.com/binaries/nomad/docs/TDS_Nomad_data_sheet.pdf) for more features.

- Personal use
- Raw data

#### **Why use this tool?**

- Allows field workers to collect and record data directly, use current taxonomically correct checklists, take low resolution images, and record sounds in a rugged handheld computer.

#### **Who will use this tool?**

- Data creation
  - Amateurs
  - Experts
  - Interest groups
- Data providers
  - Private collections
  - Casual users
- Some skills required

#### **How will the tool be used?**

- Stand-alone hand held computer
- Wireless 802.11g, Bluetooth, USB, RS-232 option
- Microsoft Office Mobile – MS Word, Excel, PowerPoint, Outlook
- Internet Explorer Mobile
- Windows Media Player
- User input required
- Accessories: cases, batteries, memory, GPS cards, etc

#### **Where in the data chain could this tool be used?**

- Data source

#### **When could this tool be used?**

- At time of data collection
- Before data is made available to ALA
- As data is imported into ALA for storage

#### **Availability**

- TDS Tripod Data Systems: <http://www.tdsway.com/>
- TDS Nomad brochure: [http://www.tdsway.com/binaries/nomad/docs/TDS\\_Nomad\\_data\\_sheet.pdf](http://www.tdsway.com/binaries/nomad/docs/TDS_Nomad_data_sheet.pdf)
- Australian distributor: Gaia Resources <http://www.gaiaresources.com.au/>
- Cost: About \$3000 plus accessories<sup>559</sup>

#### **Comments**

This is one example of a portable field data tool. I suspect that this technology is still to mature. Portable electronic devices have been delivering improved capabilities and services over the recent past, eg GPS units, phones, PDAs, digital cameras. I would expect that we will see advances in the technical capabilities and function of these portable field devices over the coming years.

---

<sup>559</sup> Piers Higgs, Gaia Resources



It is possible that these devices may want to connect directly or indirectly to the ALA: either as data suppliers – field data, observational data; or more conventionally as a data user – for example, to download a list of butterfly species found in a certain area, complete with identification notes. Access to the ALA for PDA and other portable units will need to be considered, as these have special requirements, eg screen size, minimal keyboards.

## Collaboration

## Scratchpads

Summary	
Type of tool	Group of tools
Function	Taxonomic collaboration
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Under active development. Recent workshop April 2007
Time of use	Before data is made available to ALA
Licence	Creative commons non-commercial share-alike

**Scratchpads** are part of the EDIT Unifying Web Revisions Tools, a collection of resources for collaborative taxonomy.

### Description

**EDIT** – European Distributed Institute of Taxonomy is providing the tools and resources to make it easier for taxonomists to produce revisions. These works are available for online collaboration and enable the taxonomic community to formulate the best currently available classification for their group.<sup>560</sup>

A **revision** is meant to be a work that takes a synoptic view of some taxonomic group, possibly within a defined geographical region, examines each of the taxa involved and lays out a statement of their relationships. There is no implication of scale (number of taxa covered) or scope (range of coverage).<sup>561</sup>

**Scratchpads** are web sites that are available to any taxonomic community. EDIT supports the site and provides a variety of tools, but its design and content are managed by the taxonomic community itself. The level of IT skills necessary to achieve this has been minimised as far as possible, so no specific skills in, for example, HTML are needed. Domain names are chosen by each community.<sup>562</sup>

Scratchpads provide a community of researchers a space to work on the web. They are easy to use, adaptable to reflect the community needs, and provide powerful tools for managing data.<sup>563</sup> Scratchpads are based on the content management system Drupal.<sup>564</sup>

Drupal is a free software package that allows an individual or a community of users to easily publish, manage and organize a wide variety of content on a website.<sup>565</sup>

---

<sup>560</sup> <http://www.editwebrevisions.info/node/10>

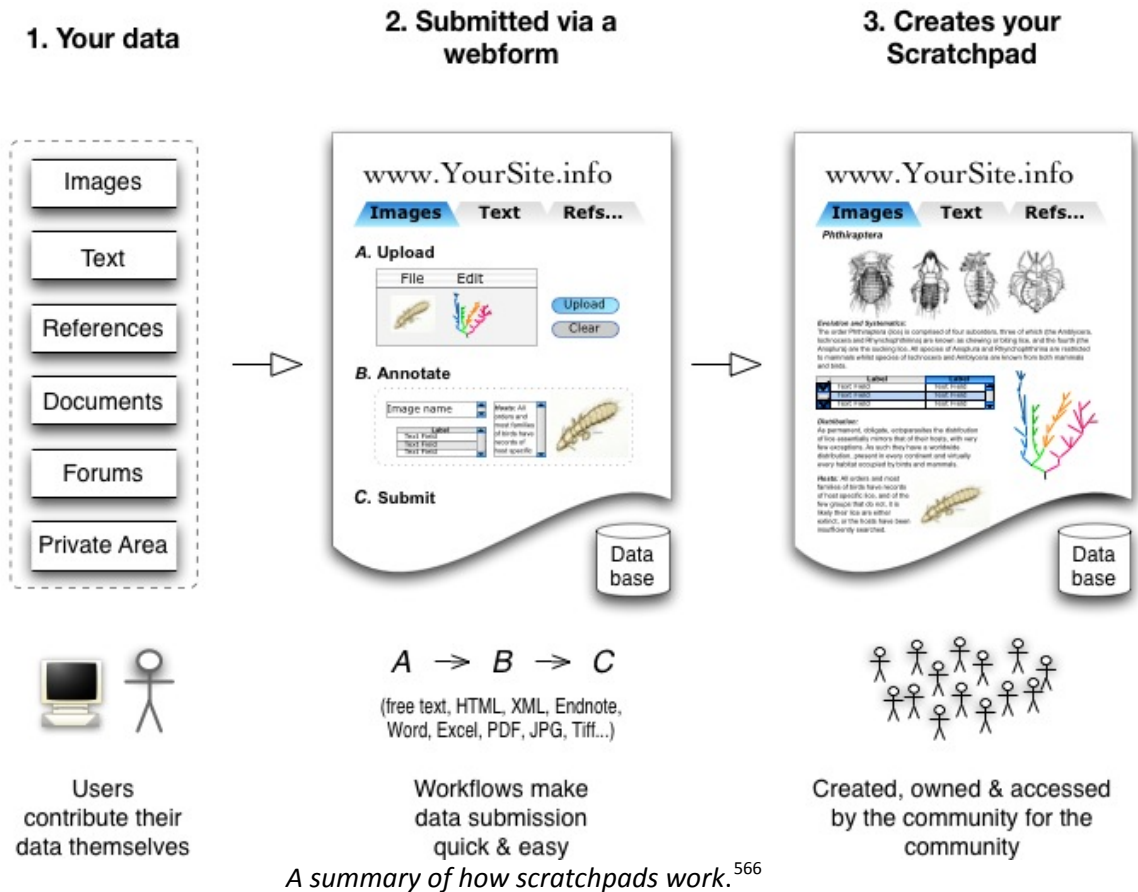
<sup>561</sup> <http://www.editwebrevisions.info/node/10>

<sup>562</sup> <http://www.editwebrevisions.info/node/10>

<sup>563</sup> <http://www.editwebrevisions.info/scratchpads>

<sup>564</sup> <http://www.editwebrevisions.info/aboutus>

<sup>565</sup> <http://drupal.org/about>



### Function

- Taxonomy
  - Identification tools, keys
- Provider interaction
  - Data preparation
  - Feedback
  - Communication
- User interface
  - Personal use
  - Visual presentation

### Why use this tool?

- For web-based community collaboration

### Who will use this tool?

- Data creation
  - Experts - taxonomy
- Taxonomy skills essential, web skills minimal

### How will the tool be used?

<sup>566</sup> <http://www.editwebrevisions.info/scratchpads>

To host Drupal:

- Drupal is written in PHP, uses MySQL
- EDIT version of Drupal uses Linux and Apache web server<sup>567</sup>
  
- Scratchpad is an online tool
- Browser required
- User input required

**Where in the data chain could this tool be used?**

- Data source

**When could this tool be used?**

- Before data is made available to ALA

**Availability**

- EDIT Work Package 6 – Unifying Web Revisions: <http://www.editwebrevisions.info/>
- Scratchpads: <http://www.editwebrevisions.info/scratchpads>
- Scratchpad FAQ: <http://www.editwebrevisions.info/aboutus>
- Drupal: <http://drupal.org/>
- Cost: free
- Conditions of Use: <http://www.editwebrevisions.info/conditions>
- Licence: Public content must be made available to users under a Creative Commons “Attribution-NonCommercial-ShareAlike” licence.<sup>568</sup>

**Comments**

The ALA could host a scratchpad-type service locally, as a service to specialists working on local collaborative projects.

---

<sup>567</sup> <http://www.editwebrevisions.info/aboutus>

<sup>568</sup> <http://www.editwebrevisions.info/aboutus>

## Other

## BioNet

Summary	
Type of tool	Application
Function	Mapping specimens
Online / Desktop	Online
Computer infrastructure	Web browser
Development status	Operational
Time of use	At the time of user request
Licence	

**BioNet** is a tool for mapping the source of flora and fauna specimens held by several NSW government agencies.

### Description

BioNet allows access to the biodiversity database of the Australian Museum, NSW Department of Environment and Conservation, and NSW Department of Primary Industries.

The system allows for users to search either for records of specific species, or for all species across a specific area of NSW. The results of a search can then be viewed either in the form of a listing, or as an interactive map.

Data used by the system is regularly updated by the participating NSW Government Agencies, and when uploaded for use on this website is checked for validity and consistency.

A key component of BioNet is the use of Master Names Lists for Taxonomy, which are utilised to validate the taxonomic Scientific Names for species data held by the participating agencies.<sup>569</sup>

Searches can be limited to species, threatened or protected.

### Function

- Analysis tools
  - Simple – distribution
- Visualisation tools
  - Maps

### Why use this tool?

- To map the source of specimen species

### Who will use this tool?

- Data users
  - Expert
  - General public

### How will the tool be used?

- Data requirements: Institution specimen data
- Online web browser
- User input required

---

<sup>569</sup> <http://www.bionet.nsw.gov.au/UnderstandingBioNet.cfm>

**Where in the data chain could this tool be used?**

- User's machine

**When could this tool be used?**

- At the time of a user request

**Availability**

- Contact: <http://www.bionet.nsw.gov.au/>
- Owner: NSW Government
- Cost: free

**Comments**

- NSW only



## DIVA-GIS

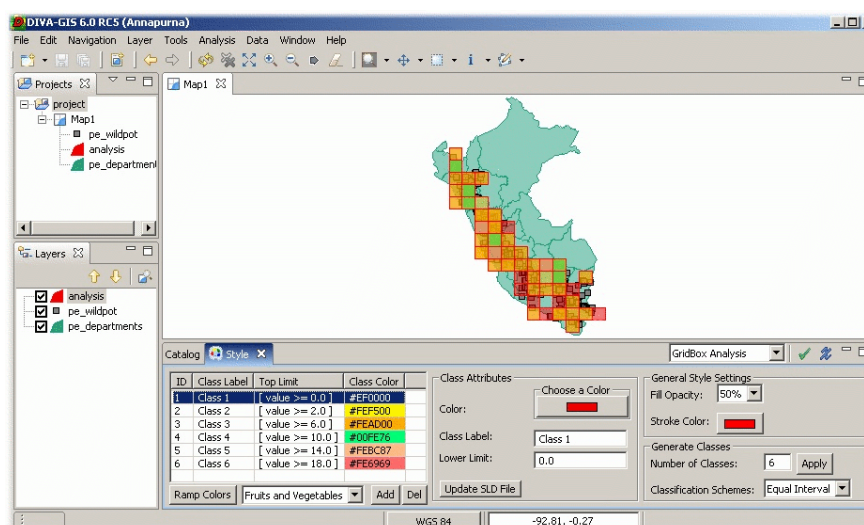
Summary	
Type of tool	Application
Function	GIS mapping, biodiversity analysis
Online / Desktop	Desktop
Computer infrastructure	Windows, Linux
Development status	Active, Version 6.0.3, January 2007
Time of use	As a post process, after data is with the user
Licence	GNU General Public License (GPL)

**DIVA-GIS** is a free geographic information system program. It is particularly suited for mapping and analysing biodiversity data, such as the distribution of species, or other point-distributions.<sup>570</sup>

### Description

With DIVA-GIS you can:<sup>571</sup>

- Make maps of the world or of a very small place, integrating, for example, state boundaries, rivers, a satellite image, and the locations of sites where an animal species was observed.
- Make grid maps of the distribution of biological diversity, to identify “hotspots” and areas that have complementary levels of diversity.
- Map and query climate data. Predict species distributions using the BIOCLIM or DOMAIN models. Create ESRI shapefiles, import and export grid data, and much more!



*Analysis richness demo.*<sup>572</sup>

DIVA-GIS is highly compatible with other GIS programs such as ArcView and IDRISI. It uses ArcView shapefiles and it uses gridfiles that can be exported to and imported from standard grid data formats.<sup>573</sup>

DIVA-GIS has been developed at CIP (International Potato Center, Peru).<sup>574 575</sup>

<sup>570</sup> <http://www.diva-gis.org/>

<sup>571</sup> <http://www.diva-gis.org/>

<sup>572</sup> <http://research.cip.cgiar.org/confluence/display/divagis/Screenshots>

<sup>573</sup> <http://www.diva-gis.org/faq.htm>

<sup>574</sup> <http://research.cip.cgiar.org/confluence/display/divagis/Home>

## Function

- Visualisation tools
  - Maps
- Analysis tools
- Non-biological data
  - Environmental data
- User interface
  - Personal use
  - Visual presentation

## Why use this tool?

- A free GIS tool with an emphasis on biodiversity analysis.
- Commercial alternatives are available. eg ESRI ArcGIS, MapInfo

## Who will use this tool?

- Data users
  - Expert
  - Interest groups
- Special skills required

## How will the tool be used?

- Compatible data available: <sup>576</sup>
  - Country level - roads, boundaries, gazetteers
  - Global climate – WorldClim, CCM3 future climate model data
  - Species occurrence data – GBIF, OBIS, ORNIS, ...
  - Satellite images - Landsat
- Desktop application
- Windows or Linux
- User input is required

## Where in the data chain could this tool be used?

- User's machine

## When could this tool be used?

- As a post process, after data is with the user

## Availability

Alternative versions are available

DIVA-GIS Version 5.2 for Windows (Sep 2005)

- DIVA-GIS: <http://www.diva-gis.org/>
- Command line version AVID-GIS: <http://www.diva-gis.org/avid.htm>
- Download: <http://www.diva-gis.org/download.asp> registration is voluntary
- Documentation: <http://www.diva-gis.org/Materials.htm>
- Manual: [http://www.diva-gis.org/docs/DIVA-GIS5\\_manual.pdf](http://www.diva-gis.org/docs/DIVA-GIS5_manual.pdf)
- Tutorial: [http://www.diva-gis.org/docs/DIVA-5\\_Tutorial.pdf](http://www.diva-gis.org/docs/DIVA-5_Tutorial.pdf)
- FAQ: <http://www.diva-gis.org/faq.htm>

---

<sup>575</sup> <http://research.cip.cgiar.org/confluence/display/divagis/Acknowledgements>

<sup>576</sup> <http://www.diva-gis.org/Data.htm>

DIVA-GIS Version 6 (Annapurna) Jan 2007

- DIVA-GIS: <http://research.cip.cgiar.org/confluence/display/divagis/Home>
  - GCP DIVA-GIS: <https://cropforge.org/projects/gcpdivagis/>
  - Download: <http://research.cip.cgiar.org/confluence/display/divagis/Downloads>
  - Manual: [https://cropforge.org/frs/download.php/276/Manual\\_Diva.pdf](https://cropforge.org/frs/download.php/276/Manual_Diva.pdf)
  - News: [https://cropforge.org/news/?group\\_id=34](https://cropforge.org/news/?group_id=34)
- 
- Cost: free
  - Licence: GNU General Public License (GPL)

## Comments

## ESRI Explorers

Summary	
Type of tool	Applications
Function	GIS viewers
Online / Desktop	Desktop
Computer infrastructure	Windows, Mac, Linux
Development status	Commercial and current
Time of use	At time of user request, as a post-process, when data is with user
Licence	Restrictive and specific for each product and application

Freely available geographic data viewers enabling institutions to provide access to GIS.

### Description



ESRI have released three products to enable free access to GIS:<sup>578</sup>

ArcGIS Explorer is a lightweight desktop client for ArcGIS Server. You can use it to connect via the Web to ESRI-hosted servers that offer unlimited access to ArcGIS Online services. You can also use ArcGIS Explorer to combine ArcGIS Online services with your local data (e.g., shapefiles, geodatabases, ArcGIS Server, ArcIMS, KML, JPEG 2000, GeoTIFF, MrSID, IMG) and other Web services. ArcGIS Explorer lets you perform GIS analysis using tasks (e.g., visibility, modelling, and proximity search).

ArcWeb Explorer demonstrates the types of rich media applications that can be developed using the ArcWeb Services JavaScript API and Adobe Flash technology. ArcWeb Explorer is not a download but instead runs as a Web application within a Web browser. It renders maps quickly in the browser instead of creating maps on the server and transmitting them to the browser. In addition to mapping streets data and imagery, ArcWeb Explorer exposes other ArcWeb Services through various widgets.

ArcExplorer is a lightweight GIS data viewer written in Java that is used to perform basic GIS functions (e.g., view, navigate, and query). It is a downloadable application that operates in a stand-alone environment and does not need to connect to a server. ArcExplorer is heavily used in the education world and runs on both Windows and Apple Macintosh computers.

### Function

- Visualisation tools
  - Maps
- Non-biological data

<sup>577</sup> <http://www.esri.com/software/arcexplorer/explorer.html>

<sup>578</sup> <http://www.esri.com/software/arcexplorer/explorer.html>

- Environmental data
- User interface
  - Personal
  - Visual presentation

#### **Why use this tool?**

- To view GIS data

#### **Who will use this tool?**

- Data providers
  - Institutions
- Data users
  - Expert
  - Interest groups
  - General public
- Limited GIS skills required

#### **How will the tool be used?**

ArcGIS Explorer

- Desktop
- Client to ArcGIS Server

ArcWeb Explorer

- Online web application running within a browser
- Renders maps supplied through ArcWeb Services

ArcExplorer

- Desktop
- Java, Windows and Mac

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- At the time of a user request
- As a post process, after data is with the user

#### **Availability**

- ESRI Explorers: <http://www.esri.com/software/arcexplorer/explorer.html>
- ArcGIS Explorer: <http://www.esri.com/software/arcgis/explorer/index.html>
- ArcWeb Explorer: <http://www.arcwebservices.com/awx/index.jsp>
- ArcExplorer: <http://www.esri.com/software/arcexplorer/index1.html>
- Cost: free
- Licence: Restrictive and specific for each product and application  
see <http://www.esri.com/legal/licensing/software-license.html>

#### **Comments**

As much of the data within the ALA is georeferenced, being able to present this and other geographically based data, in a form that can be viewed is important. These Explorer products will enable a wider reach for well-presented geospatial information.

## GBIF RESTful Web Services

Summary	
Type of tool	Web services portal
Function	Data transfer
Online / Desktop	Online
Computer infrastructure	Direct access to XML data
Development status	Active
Time of use	At time of user request
Licence	GBIF Data Use agreement – respect, attribution

The **GBIF portal** offers a range of services to support the use of biodiversity data.<sup>579</sup>

### Description

The portal includes a range of web services that can be used by other portals and applications to directly access XML formatted GBIF data. The services currently available include:<sup>580</sup>

- **Taxon data service** - This service (<http://data.gbif.org/ws/rest/taxon>) provides a range of options for viewing information on the names and classifications used by the different datasets that are accessible through the GBIF portal. Data are returned using the **TDWG Taxon Concept Schema**. See the [documentation](#).
- **Occurrence record data service** - This service (<http://data.gbif.org/ws/rest/occurrence>) provides a range of filters for selecting occurrence records. The currently supported response formats include **TDWG Darwin Core** records and **KML** (for use with **Google Earth**). See the [documentation](#).
- **Occurrence density data service** - This service (<http://data.gbif.org/ws/rest/density>) returns summary counts of occurrence records by one-degree cell for a single taxon, country, dataset, data provider or data network. The currently supported response formats include a simple XML format and **KML** (for use with **Google Earth**). See the [documentation](#).
- **Dataset metadata service** - This service (<http://data.gbif.org/ws/rest/resource>) returns metadata relating to the datasets accessible through the GBIF portal. The currently supported response format is a simple XML structure. See the [documentation](#).
- **Data provider metadata service** - This service (<http://data.gbif.org/ws/rest/provider>) returns metadata on the data providers sharing data through the GBIF portal. The currently supported response format is a simple XML structure. See the [documentation](#).
- **Data network metadata service** - This service (<http://data.gbif.org/ws/rest/network>) returns metadata on the data networks connected to the GBIF portal. The currently supported response format is a simple XML structure. See the [documentation](#).

### Function

- Taxonomy
- Provider interaction
  - Communication
- Metadata
  - Data discovery
- Interface
  - Computer use
  - Raw data

<sup>579</sup> <http://data.gbif.org/tutorial/services>

<sup>580</sup> <http://data.gbif.org/tutorial/services>

### Why use this tool?

- Direct machine access to data

### Who will use this tool?

- Data creation
- Data capture
- Data providers
  - Institutions
  - Private collections
- Data users
  - Expert
  - Interest groups
- ALA infrastructure
- Special skills required

### How will the tool be used?

This is a collection of RESTful web services to support data access

#### Data requirements

- TDWG Taxon Concept Schema – for Taxon data service
- TDWG Darwin Core and KML – for Occurrence record data service
- XML – for Occurrence density data service and metadata services
- KML – for Occurrence record data service and Occurrence density data service
- Online web services for computer interaction
- User input is generally one layer removed
- Ideal for batch operation

### Where in the data chain could this tool be used?

- Data source
- ALA central
- User's machine
- Pathways between these

### When could this tool be used?

- Before data is made available to ALA
- As data is imported into ALA for storage
- At the time of a user request

### Availability

- GBIF RESTful services: <http://data.gbif.org/tutorial/services>
- Documentation:
  - Taxon data service: <http://data.gbif.org/ws/rest/taxon>
  - Occurrence record data service: <http://data.gbif.org/ws/rest/occurrence>
  - Occurrence density data service: <http://data.gbif.org/ws/rest/density>
  - Dataset metadata service: <http://data.gbif.org/ws/rest/resource>
  - Data provider metadata service: <http://data.gbif.org/ws/rest/provider>
  - Data network metadata service: <http://data.gbif.org/ws/rest/network>
- Cost: free

- GBIF Data Use Agreement – a framework of due attribution:  
<http://data.gbif.org/tutorial/datauseagreement>

#### Comments

- The KML map layers can be used, for example, with Google Earth.
- A clear introduction to REST: [How I Explained REST to My Wife](#) by Ryan Tomayko.<sup>581</sup>

---

<sup>581</sup> <http://naeblis.cx/articles/2004/12/12/rest-to-my-wife>



# HerpNet

Summary	
Type of tool	Federation of databases
Function	Reptile and amphibian taxonomy
Online / Desktop	Online
Computer infrastructure	Browser
Development status	Active development
Time of use	
Licence	Research. Not for redistribution.

**HerpNet** is a collaborative effort by natural history museums to establish a global network of herpetological collections data.

## Description

The mission of HerpNet is to bring the accumulated knowledge from more than four million specimens in world-wide museum collections into currency for science and society by creating a distributed database with access from various portals. HerpNet will connect large repositories of information with smaller collections that have regional specializations. Similar efforts (e.g. **MaNIS**, **FishNetII**, **MaPSTeDi**, **ORNIS**) are being accomplished for other taxa and regions and the herpetological community is poised to make its own contribution to the study of biodiversity.

HerpNet has included:

- Species query<sup>582</sup>
- Gazetteer<sup>583</sup>
- Georeferencing resources<sup>584</sup>
- HerpNet email list<sup>585</sup>

Currently, 59 institutions are participating in the HerpNet community, with an open-ended invitation to institutions who would like to join.

## Function

- Visualisation tools
  - Maps
- User interface
  - Personal use
  - Raw data and visual presentation

## Why use this tool?

- HerpNet allows access to data associated with specimens that may be collected in Australia, but are held by external institutions.

## Who will use this tool?

- Data capture
  - Curators

---

<sup>582</sup> <http://www.herpNet.org/herpNet/portal.html>

<sup>583</sup> <http://www.herpNet.org/herpNet/Gazetteer/search.html>

<sup>584</sup> <http://herpNet.org/Gazetteer/GeorefResources.htm>

<sup>585</sup> <http://www.herpNet.org/herpNet/communication.html>

- Data users
  - Expert
  - Interest groups

#### **How will the tool be used?**

- Online databases
- Mapped results are displayed through BerkeleyMapper
- User input is required

#### **Where in the data chain could this tool be used?**

- User's machine

#### **When could this tool be used?**

- At the time of a user request

#### **Availability**

- HerpNet: <http://www.herpnet.org/>
- Search: <http://www.herpnet.org/portal.html>
- Resources: <http://www.herpnet.org/Gazetteer/GeorefResources.htm>
- Gazetteer: <http://www.herpnet.org/herpnet/Gazetteer/search.html>
- Server specifications: <http://www.herpnet.org/herpnet/documents/severspecs.html>
- Listserv: <http://www.herpnet.org/herpnet/communication.html>

#### **Comments**

## OZCAM – Online Zoological Collections of Australian Museums

Summary	
Type of tool	Distributed databases
Function	Taxonomy, species distribution
Online / Desktop	Online
Computer infrastructure	Browser, HTML, CSV, Text, XML
Development status	Experimental
Time of use	Before data is made available to ALA. At time of user request
Licence	Research and personal use only

**OZCAM** - Online Zoological Collections of Australian Museums - provides an online distributed query network to faunal collections in Australian museums.<sup>586</sup>

### Description

Currently OZCAM provides a "proof of concept" that it is possible to provide access to datasets around Australia via a single portal. Currently, queries can be made on five groups of animals: mammals, reptiles and amphibians, fishes, dung beetles and molluscs. An extensive query screen allows the user to query by taxon name, by place, by time or by a combination of these. The user can also nominate which museums are to be queried. Data returned from sites around Australia can be displayed either as dots on a map or as a list. From either of these displays it is possible to mine the data further and view the raw dataset for any single specimen.<sup>587</sup>

Australian Virtual Herbarium

OZCAM has been designed to parallel the Australian Virtual Herbarium, a site where queries can be made about the flora of Australia.

OZCAM is an initiative of the Council of Heads of Australian Faunal Collections (CHAFC), a body formed from representatives of every Australian museum that holds significant faunal collections. Development of OZCAM has been funded through contributions by participating institutions and the Australian Government, through the Department of the Environment and Heritage.<sup>588</sup>

### Function

- Visualisation tools
  - Maps
- Taxonomy
- User interface
  - Personal and institutional use
  - Raw data and visual presentation

### Why use this tool?

- View holdings in Australian collections

### Who will use this tool?

- Data creation

---

<sup>586</sup> <http://www.ozcam.gov.au/about.php>

<sup>587</sup> <http://www.ozcam.gov.au/about.php>

<sup>588</sup> <http://www.ozcam.gov.au/about.php>

- Experts - taxonomy
- Data capture
  - Curators
- Data providers
  - Institutions
- Data users
  - Expert
  - Interest groups

#### **How will the tool be used?**

OZCAM has been designed to demonstrate a "proof of concept" that faunal collections around Australia can be queried simultaneously on a distributed network via a single portal. The site is powered by software from KE Software and hosted on a server run by Museum Victoria.

Each institution holds data on its own server in whatever form the institution chooses. OZCAM has not imposed a standard database but has simply requested data in a specific number of fields. Different museums have chosen different modes of delivery of their data: some have published datasets outside their network's firewall; others allow direct access to the datasets through their security systems.

Communication between the distributed datasets and the central portal is achieved by using wrapper software developed by AMOL and Museum Victoria, to interpret queries and responses. An XML schema gives structure to the datasets.<sup>589</sup>

- Formats are HTML, CSV, Text, XML
- Online search
- User input required

#### **Where in the data chain could this tool be used?**

- Data source
- User's machine

#### **When could this tool be used?**

- Before data is made available to ALA
- At the time of a user request

#### **Availability**

- OZCAM: <http://www.ozcam.gov.au/>
- Search: <http://www.ozcam.gov.au/cgi-bin/emu-dataportal.cgi>
- About OZCAM: <http://www.ozcam.gov.au/about.php>
- Copyright: Research and personal use only. See: <http://www.ozcam.gov.au/copyright.php>

#### **Comments**

The ALA may use some of the concepts of OZCAM.

---

<sup>589</sup> <http://www.ozcam.gov.au/about.php>

## Spotter – SPIRE Ontology Tool

Summary	
Type of tool	Application
Function	Observational data creation
Online / Desktop	Online
Computer infrastructure	Firefox extension
Development status	Experimental
Time of use	Before data is made available to the ALA
Licence	Non-commercial, personal use

**Spotter** is a Firefox RDF-based extension designed for observations made by citizen scientists from unstructured sightings of organisms (e.g., in web blog entries, discussions, photo-sharing sites).

### Description

The screenshot shows a web browser window with the URL <http://spire.umbc.edu>. The page title is "Observation Form". The main content is a satellite map of a coastal region with a red location pin. Below the map, there are input fields for "Date" (02 mm 26 dd 2008 yyyy), "Common Name" (Fairy penguin), and "Scientific Name" (Eudyptula minor). A "Find Scientific Name" link is also visible.

*Spotter observation form.*

**Spotter** allows you to easily create RDF/OWL for species observations. It places a button on your Firefox toolbar that calls up a form with fields taken from our observation ontology (Observer, Reporter, Location, Common Name, Scientific Name, etc.). Upon submitting the form, an OWL observation record is created and stored on our server. You are presented with a link to the record, which you can then link to from your blog.<sup>590</sup>

<sup>590</sup> <http://spire.umbc.edu/firefox/>

By following the link, semantic web crawlers then generate and index the appropriate RDF.<sup>591</sup>

Spire, based at UMBC, University of Maryland Baltimore County, is a research project investigating how semantic web technologies can be used to support science in general and the field of ecoinformatics in particular.<sup>592</sup>

### Function

- Provider interaction
  - Observational data
- User interface
  - Personal use
  - Raw data

### Why use this tool?

- Spotter has been designed to make it easier for individual scientists to convert their information to RDF and OWL.<sup>593</sup>
- Individual observations could be directly entered into a database within the ALA

### Who will use this tool?

- Data creation
  - Non-experts
- Data providers
  - Casual users
- Semi-skilled expertise

### How will the tool be used?

Spotter is a Firefox extension.

Once Spotter is installed, you'll notice that there are a few changes to your browser UI: Under "Tools", you'll see (1) "Spotter Options..." and (2) "Report a nature observation with Spotter".

1. is where the UI for setting default values for reporter, observer, etc. are
2. is where you can launch the spotter form and submit an observation.

To quickly access the default value preferences, you can click on the little spotted owl at the bottom right of your browser window.<sup>594</sup>

- Desktop tool
- Firefox web browser
- User input required

### Where in the data chain could this tool be used?

- Data source

### When could this tool be used?

---

<sup>591</sup> <http://www.tdwg.org/proceedings/article/view/270>

<sup>592</sup> <http://spire.umbc.edu/us/>

<sup>593</sup> <http://www.tdwg.org/proceedings/article/view/270>

<sup>594</sup> <http://spire.umbc.edu/firefox/>

- As data is made available to ALA

**Availability**

- Spotter: <http://spire.umbc.edu/firefox/>
- SPIRE: <http://spire.umbc.edu/>
- Experimental
- Licence: non-commercial, personal use <http://spire.umbc.edu/legal/>

**Comments**

- Spotter delivers data to the SPIRE research group.
- Spotter is an experimental application and is an example of remote observational data entry. The ALA could use a modified form of this to enable direct entry of observational data.