Fauna collections databases in the Western Australian Museum

P F Berry & G Christie

Western Australian Museum
Francis Street, Perth WA 6000
email: paddy.berry@museum.wa.gov.au

Abstract

The history of computerisation of the Western Australian Museum's modern fauna collections is outlined (it was initiated in the late 1970's on the Treasury mainframe computer). The Museum is custodian of approximately 1.5 million specimens and specimen lots of modern fauna. The majority of specimens in its collections are from Western Australia, but the collections also include specimens from Indonesia and New Guinea. Currently only the vertebrates are comprehensively databased. The Museum's current primary collections database system is briefly described as well as its recently released on-line version – *FaunaBase*.

Keywords: Database, museum, Faunalist, FaunaBase

Historical background

The first faunal specimens with associated data were accessioned into the Western Australian Museum collection in the 1890's. As was traditional with natural history museums, specimens were (and still are) arranged in collections corresponding to phylogenetic groupings (mammals, birds, arachnids, *etc*). The practice of assigning each specimen within a collection a unique registration number, recording associated data on a label, and entering duplicate and additional data into a hand written catalogue was started in the early 1900's. However, different curators adopted somewhat different methods of assigning registration numbers. In some collections the specimens were numbered sequentially, whereas in others they were numbered sequentially by year. Also, data accompanying many of the earliest records are often imprecise, particularly the collection locality (*e.g.* "Western Australia").

The first attempt to produce an electronic listing (but not a database in the modern sense) was in the late 1970's when a catalogue of the modern mammal collection was produced using the Treasury mainframe computer and involving punch cards and special programming. In 1983, as a foundation member of the Western Australia Land Information System, the Museum received funding for an in-house UNIX server through the Government Computing Committee and the Museum employed a contract programmer (with a background in biology) to develop customised database software for the mammal, herptile, fish, arachnid and crustacean collections. This in-house system ran very successfully until the late 1980's when the Museum's inability to continue to support it prompted a switch to a commercial multi-user database system (INGRES), developed and supported externally. The Museum's collections databases were temporarily transferred to FOXPRO in about 1995, pending availability of BIOLINK, a multi-user database designed for biological collections management and research under development by CSIRO.

BIOLINK was not designed with registration number as a key field. For this and other reasons the Museum decided in 2000 to adapt the collections management system developed in Microsoft Access® by the Department of Agriculture for its entomology collection. This was done by Dr Graeme Christie and is the system currently in use and described below.

The Museum's primary collections databases

The approach taken was to develop a generic database system that accommodated the collections management requirements of all the Museum's fauna collections. This necessitated considerable effort to standardise fields and existing differences in data standards before all collections could be successfully migrated to the new system. Post migration required extensive validation of data.

Currently there are eleven faunal collections databases operating on this generic Microsoft Access® system. The

Table 1

List of modern fauna collections in the Western Australian Museum, numbers of records databased and proportion of each collection databased.

Collections	Numbers of databased records	Approximate percentages of collections databased
Mammals	49,734	100
Sub-fossil Mammals	10,000	15
Birds	33,822	100
Reptiles / Amphibians	131,847	100
Fishes	47,748	100
Arachnids	46,666	<10
Insects	34,183	<10
Molluscs	22,400	<10
Crustaceans	31,056	30
Worms	4,277	<10
Other Marine Invertebrates	12,062	<10

© Royal Society of Western Australia 2002

This paper was presented in similar form at the Workshop on the Co-ordination and Integration of Terrestrial Vertebrate Fauna Survey Database for Western Australia, held on April 5, 2002, at the University of Western Australia

size of each, and degree to which data capture of each collection is complete varies and is summarised in Table 1. The vertebrate collections are the most comprehensively databased, the invertebrate groups are the least (because of the larger numbers of taxa and poorer knowledge of the taxonomies).

The Museum's primary collections databases were designed with ease-of-use being the primary aim without losing the potential to expand or accommodate new technology as it became available. Being based on Microsoft Access®, the database technology follows what has become an industry standard and as such is well understood and supported; an important consideration for ongoing development and maintenance. This technology also allows for immediate integration to a wide spectrum of other existing or newly developed computer technologies, in particular the Internet.

The ability to rapidly and accurately enter large quantities of information into the primary databases has been addressed in the database design. Predefined forms are used to enter the data for one record at a time: each record is uniquely identified by its Registration Number, which doubles as the collection Lot number. Wherever feasible, for example, in the selection of taxonomic names, country, state, etc, look-up lists are provided for ease of data entry and integrity. Data validation is enforced at field level to avoid the entry of invalid or inaccurate data. Automated processes, such as the ability to automatically derive higher taxa from genus, or the selection of default values, have been implemented. An interface to an external gazetteer database provides the ability to rapidly obtain latitudes and longitudes from selected place names and / or offsets from these places, or *vice versa*. Automatic conversion from latitude to decimal is also implemented.

An interface to a 'Tissues' database is provided in order to cater for the increasing need to use molecular technology to support taxonomy. Similarly, an interface to an Images database links specimen records to digitised images.

A suite of utilities has been developed for rapid database maintenance. In order to ensure that collections database records align with taxonomic changes a program has been developed using Visual Basic. This program allows global changes to the database records whenever changes to taxonomic names are required or if taxa are incorrectly positioned within the taxonomic hierarchy. A taxonomic tree is generated based on the taxa within the database and any taxon can be moved to its correct location in the hierarchy using a "drag-and-drop" method. All SQL statements required to implement these global changes are automatically generated by the program. This utility can also generate a distribution map for any taxon. Each specimen is represented by a single dot on the map indicating its collection location. Underlying records for that taxon can then be interrogated and edited, if required, by selecting the dot or dots. A related program provides a read-only taxonomic tree / mapping interface. Another utility allows up to 20 related or unrelated taxa to be plotted superimposed on the same map for comparison. This program is useful for rapidly finding and editing records with incorrect geographical coordinates.Mıan thepidlaIn o Undemliomı