

# Overview of the CEGSA digital database focusing on the Articles data set

Graham Pilkington

Cave Exploration Group of South Australia

The CEGSA digital database is an attempt to have all of our data available at call. It has a hierarchical index overview into relational subsets. The current data covers library items such as journal and book holdings, management items such as membership and meetings, and karst records information.

To enable many people to contribute, the database is grouped into subsets, each of which can operate as an independent database.

The articles database handles references to articles of all types. Currently this is limited to newspapers and magazines, caving journals, and trip reports, but it's intended to include any document that can be referenced. Each type of reference data is itself a subset of the Articles database, and can be operated independently, but needs a special merge to realign into the main dataset. In keeping with the CEGSA database concept, not only is there a detailed index of each article, but the actual article can be displayed.

## Preamble.

This workshop is presented to encourage other clubs to digitize their own Records and Library. It can be used as an aid in developing your own system. Copies of the software can be made available on request adjusted to your own club.

The CEGSA database began as a means of listing what we physically had in our Library and Records. It was a simple index and organized around the type on media being stored. For example, books, journals, maps, photographic slides and stills, videos, and "cave master sheets". When digitised information became more prevalent and the decision to make the actual data available in digital format, it became apparent that the media on which the data are stored is only relevant because of the quality implied.

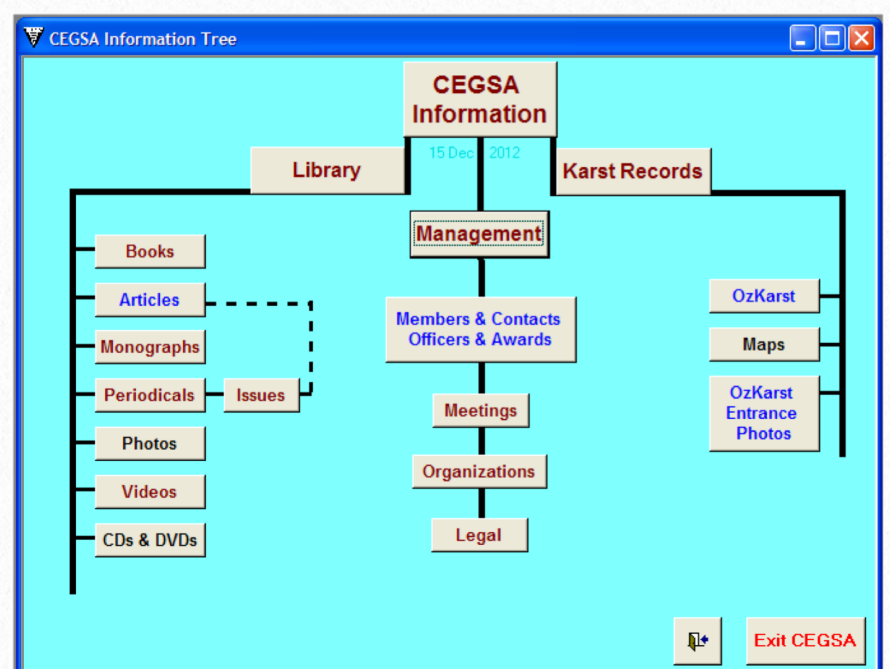
As an example, photographs were first stored as prints, negatives, and slides and needed to be indexed accordingly. Photos can also come as digital images; scanned images; included

in a report, journal or book; on hard drives, CD's, DVDs or Videos. Note the mixture of media and source. Data requests are rarely for a specific media but for the content – such as photo taken inside a specific cave. It was decided to ignore the media and convert the CEGSA database into a contents index and full retrieval system.

The following is a description of what we have – a mixture of media and content – as the database is slowly converted from a media index to content retrieval.

## The CEGSA digital database.

The CEGSA digital database is an attempt to have all of our data available at call. It has a hierarchical index overview into relational subsets. It's written in the relational database software Microsoft ACCESS. The start-up screen is as follows:



To put everything into digital format is not only too great a task for one person, but the data are required to be up-to-date for specific Club Officers. CEGSA has implemented this by having independent databases linked to a central core so that officers can control the data subset appropriate to their function in the Group. On replying to requests, the officer is only responsible for their own data, linking in any other data required from the latest data subsets that they hold.



Currently the subsets that we have are:

*Library & Records* as the central focal database. Journal, book, CD and map holdings. Currently includes group management such as meetings and minutes.

*People*, including membership, office bearers, awards, and contacts.

*OzKarst* for all karst data, where each caving area can be controlled by a different person.

*Articles*, where each article data subset can be controlled by a different person (Newspapers and Magazines; Caving Journals; Trip Reports; ..)

Our data are divided into three logical sections:

*Management items* to do with running the club,

*Karst Records* to store data pertaining to caves and karst, and

*Library* to store published material.

In the data tree shown above, buttons with brown text indicate that the data is held in the core database and looked after by the Records Keeper. Buttons with blue text indicate that the data are held in an independent database, and buttons with black text indicate that the data group is still in preparation. It should be noted that that *Articles* database will be absorbing *Monographs* and probably *Books* as well.

**Management** This displays a list of the type of data stored under Management.

**Meetings** This displays a list of meetings that have been held by CEGSA.

The availability tick boxes are for:

- is known to exist,
- an image present – typically for hand-written (sometimes illegible!) minutes,
- digital text present.

The attendance column notes how many people were present. Double-clicking in this column brings up a list of the actual attendees. The names are extracted from the linked *Members & Contacts* database.

Double-clicking in the “for” column displays the actual minutes.

A secondary window appears that allows a restriction on the type of meetings displayed. This selection is stored for future use.

**Organizations** This brings up a Form displaying a selectable organization that either has an interaction with CEGSA or from which information has been obtained. This table is being expanded and united with the organization tables in all the separate databases.

The “predec” box stores the organization from which this organization originated. Originally this was to cater for clubs that change their name but it needs to be expanded so that merged organizations such as the Southern Tasmanian Cavers can show all the source clubs.

The “parent” box stores the organization that oversees this one. For example, the State Liaison Councils of the ASF would have the ASF as the parent.

Most of the started and ceased dates are missing because little research has been done on this. Similarly with the ASF 3-letter code.

**Legal** This brings up a Form displaying a selectable legal document such as our Constitution. Because we have so few documents currently entered and it is very infrequently used, the Form is in a primitive state. For instance, you need to scroll through all the available documents to find what you want. Clicking on Display retrieves the document.

**Members & Contacts / Officers & Awards** This brings up the database maintained by our Membership Officer. Naturally it's designed to suit the way that CEGSA operates. The “publish?” tick boxes enable a very detailed selection of what information is allowed to be distributed. A full history is kept of membership status and addresses but most other fields are overwritten when updated. Note that historical data before the start date of this database has not always been



Membership and contacts page

completed for superseded or now irrelevant information; eg the ASF membership history.

Family memberships can be handled, including historical memberships. Each family membership has a “prime” member just for convenience in handling the members within the family.

Membership status can be selective to show any of membership/office/awards/qualifications. If data are available for a specific status class, then it’s highlighted in colour.

Personal details can be accessed by double-clicking the names box. This shows any change in name such as when a member gets married. Without this, it can be confusing reading such things as trip reports. [Close *Details* box then select Ann Pilkington] Because she had two names, this is noted on the main form.

**Karst Records** This displays a list of the type of data stored under Karst Records. *This section is handled by OzKarst and will be covered in the next session.*

**Library** This displays a list of the type of data stored under Library.

**Books** Our books.

**Monographs** Our monographs. Note that this will be moved into the Articles database.

## Videos

Our videos. These will be merged with DVDs, CDs, and other media into the data type. For instance, movies, stills, documents, and data.

## Periodicals

A list of the periodicals held. The overlap with the *Articles* database is being sorted out.

**Catalogued** This is a tabulation of the issues of the periodicals. Some issues are available for display.

Clicking on a list brings up a quick summary for each issue, including what is bound into specific CEGSA hard covers.

Adding a new issue is largely automated.

**Articles** This brings up the *Articles* database. It’s intended to include any document that can be referenced. Each type of reference data is a subset of the Articles database and can be operated independently but needs a special merge to realign into the main dataset. In keeping with the CEGSA database concept, not only is there a detailed index of each article, but the actual article can be displayed.

The Articles page

Fred Aslin maintains the Newspapers & Magazines subset and I look after the other sub-sets such as Caving Journals and Trip Reports. This database will eventually replace the Reference section of OzKarst.