

Demo Title: OLAC Search Engine

Presented by: Baden Hughes, Amol Kamat and Steven Bird
Department of Computer Science and Software Engineering
University of Melbourne

Software Title: OLAC Search Engine

Software URL: <http://www ldc.upenn.edu/olac/search.php>

Availability: The software is open source and available under the GNU Public License.

The sources can be obtained as a part of the OLAC Tools Suite available from <http://sourceforge.net/projects/olac/>.

Description:

The purpose of the OLAC Search Engine is to provide a powerful search facility across OLAC Archives. Modelled on features common in web search engines, the service demonstrated here allows a user to explore relevant archive holdings without requiring any knowledge of the underlying data structures. Search results are displayed based on a metadata quality ranking scheme which in turn is derived from the archive's adherence to best practice recommendations in terms of metadata expressions.

The OLAC Search Engine features a number of utilities which are common in web search engines:

- Alternative Language Names Thesaurus: the Ethnologue data is used to provide a source of alternate language names which may improve search results
- Record Matching: the search engine supports exact word matching, substring matching and phrase matching for increased precision
- Inline Syntax: searches can include inline syntax derived from OLAC metadata elements eg (creator:hale)
- Combinatorial Operators: the search engine supports the use of AND, OR, NOT, +, - operators in search strings
- External Links: the search engine provides links to other web search engines with pre-composed key word queries

Search results are displayed according to a "star rating" which indicates the degree to which an archive as a whole and the matched records comply with recommendations regarding the use of OLAC metadata elements, including the controlled vocabularies for language identification, linguistic field, linguistic data types and participant roles.

We offer the OLAC Search Engine as a service to assist end users in locating language resources of interest.