

An Online Database for Computational Paremiology

Abstract / Abstract / Povzetek

The *Global Proverbs Project* harnesses common "web 2.0" technologies such as user-generated content, open application interfaces (APIs), and folksonomies to enable new avenues in paremiological research.

While many online proverb collections exist, they often do not make use of modern database and web technologies. The best of currently available databases are large lists of proverbs (usually monolingual) indexed by theme and possibly a concordance. These sites are essentially the online equivalent of the bound proverb collections found in libraries. Through the use of current web technologies and conventions, the *Global Proverbs Project*, seeks to leverage the global internet community to build a broad multilingual database annotated with useful metadata and multiple classification schemes.

By collecting a large corpus of proverbs in multiple languages from multiple cultures, the *Global Proverbs Project* seeks to allow the application of data-mining techniques to proverbs thus enabling a new method of paremiological study. The *Global Proverbs Project* allows paremiological questions to be asked and answered via computational approaches.

The *Global Proverbs Project* is a multilingual database build upon the Django web framework and MySQL. It allows proverbs to be annotated with cultural, temporal, geographic, and thematic information. Individual proverbs may be translated into any number of other languages and proverb variations may be provided. For each proverb space is provided for the discussions of meaning and the addition of citations.

The *Global Proverbs Project* seeks to be a resource for proverbs research and is open to the internet community. It builds upon the ethos of other user built knowledge repositories like Wikipedia while providing structured data and APIs useful for the study of paremiology. Interested individuals are encouraged to contribute.