

Abs-166

Eros Zanchetta (University of Bologna)

Corpora for the masses: introducing CARCASS / CORPSE, an open-source client-server architecture for easy corpus access

This poster introduces a new corpus manager designed specifically to target language professionals and students with limited computer skills.

The Corpus Archive Search System (CARCASS) is a client that allows users to query corpora (including very large annotated ones) located on remote computers running the Corpus Server (CORPSE).

The client is a cross platform, user-friendly GUI that aims to replicate the full potential of the Corpus Query Language: queries can either be built with the help of a wizard or entered directly using the CQP syntax. The system supports positional (e.g. part-of-speech) as well as structural (e.g. text_id) attributes. Unlike similar applications, the interface does not use a web browser.

The server component – built on top of OpenCWB (<http://cwb.sourceforge.net>) – manages user access and implements a sophisticated permission mechanism capable of restricting access to specific corpora and allocating computing time to users.

The system is free and open-source software.