PyCDSL: A Programmatic Interface to Cologne Digital Sanskrit Dictionaries

Hrishikesh Terdalkar   **Arnab Bhattacharya**
hrishirt@cse.iitk.ac.in   arnabb@cse.iitk.ac.in
Dept. of Computer Science and Engineering,
Indian Institute of Technology Kanpur,
India.

PyCDSL is a Python library that provides programmer friendly interface to Cologne Digital Sanskrit Dictionaries (CDSD) (cds, 2022). The library serves as a corpus management tool to download, update and access dictionaries from CDSD. The tool provides a command line interface for ease of search and a programmable interface for using CDSD in computational linguistic projects written in Python 3.

The command line interface is provided in two modes (1) a console command (`cdsl`) and (2) an interactive REPL[1] interface. Both modes come with a rich search functionality. Users may search by key (a dictionary entry) or value (meaning provided in the dictionary) or both. The search can be performed on multiple dictionaries at the same time. The interactive REPL mode can be triggered by passing option '`-i`' to the console command '`cdsl`'. Once in this mode, a user may simply input a term to perform search. Search related settings can also be changed in this mode. The library also comes with extensive transliteration support powered by Python library `indic-transliteration`[2]. The input can be provided in any of the supported schemes and the results can be exported in a scheme of choice as well. The results can be copied to clipboard (powered by *Pyperclip*[3])

Every dictionary uses specific conventions in the meaning (value) of the dictionary entries (key). The library can be extended with custom parsers to parse the entries in the dictionaries in a custom manner.

Figure 1 shows a part of the documentation detailing usage instructions for the console command. The detailed documentation can be found at `https://pycdsl.readthedocs.io/en/latest/index.html`. The library is available on Python Package Index (PyPI[4]) at `https://pypi.org/project/PyCDSL/`. The source code can be found at `https://github.com/hrishikeshrt/PyCDSL`.

References

[cds2022] 2022. Cologne Digital Sanskrit Dictionaries. https://www.sanskrit-lexicon.uni-koeln.de. version 2.0.738, accessed on October 13, 2022.

---

[1] `https://en.wikipedia.org/wiki/Read-eval-print_loop`
[2] `https://github.com/indic-transliteration/`
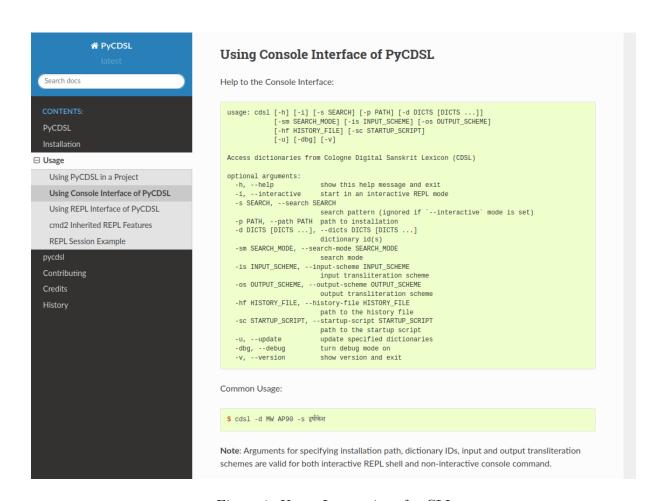[3] `https://pypi.org/project/pyperclip/`
[4] https://pypi.org/

Figure 1: Usage Instructions for CLI