



# Linguistic Data Consortium & CoreTrustSeal

Daniel Jaquette, Publications Programmer

[jaquette@ldc.upenn.edu](mailto:jaquette@ldc.upenn.edu)

- ◆ The Linguistic Data Consortium (LDC) is a non-profit, hosted at the University of Pennsylvania
  - LDC creates and distributes language data to universities, laboratories, companies and libraries for language-related education, research and technology development.
- ◆ LDC's Catalog has over 800 holdings in more than 90 languages. Three to four data sets are released monthly



- ◆ Motivation
  - Provide an international, recognize standard
  - Sanity check of catalog's quality
  
- ◆ Submission
  - We started our work during the transition of the Data Seal of Approval to CoreTrustSeal

- ◆ Some of our internal assumptions were off
- ◆ External documentation needed improvement
- ◆ Investigation into broader persistent identifiers

