Application and Adaptation of the Crow Corpus Search Interface:
An Introduction and Technical Specifications

CROW

This document provides an overview of the Corpus and Repository of Writing (Crow)'s interface. Organizations planning a web-based corpus search interface can use this document to evaluate whether Crow's open-source software can be adapted or used as a model.

This document covers:

1. Design & usability of the interface
2. A case study in adapting the software for a similar, multilingual corpus
3. Technical specifications of the software

Design & Usability: Main navigational features

- filter by draft, assignment, and country (e.g., find first drafts of synthesis papers written by Indian students)
- search by exact word or word family (e.g., "argue, argues, argued, arguing")
- See highlighted writing excerpts

Interface Design

- **Dataset:** a large-scale learner corpus of English writing samples from university foundational writing courses, as well as pedagogical materials used in those courses
- **User audience:** corpus researchers, writing teachers, and students
- **Access:** registration required; different tiers of access

---

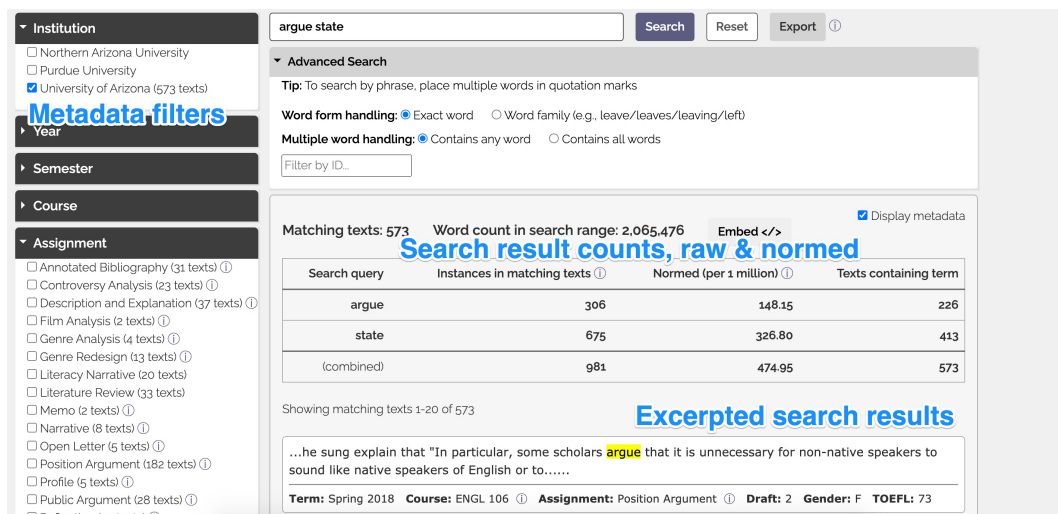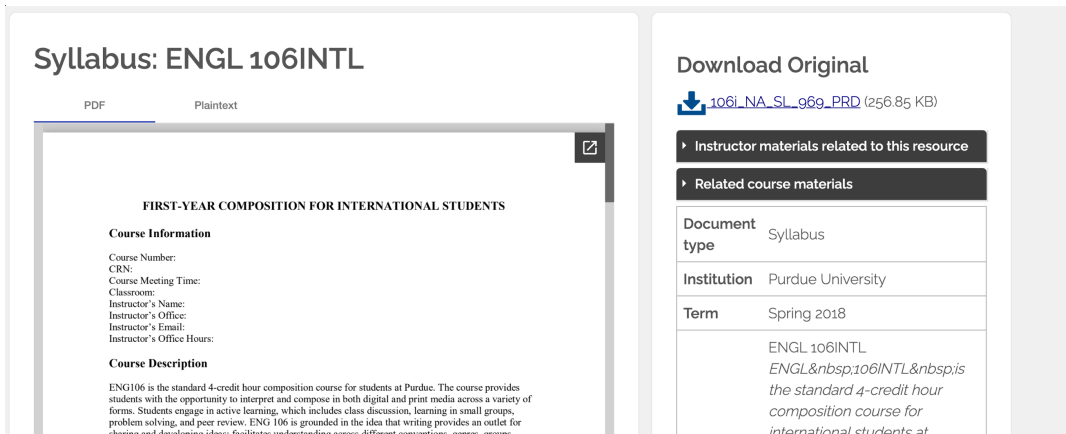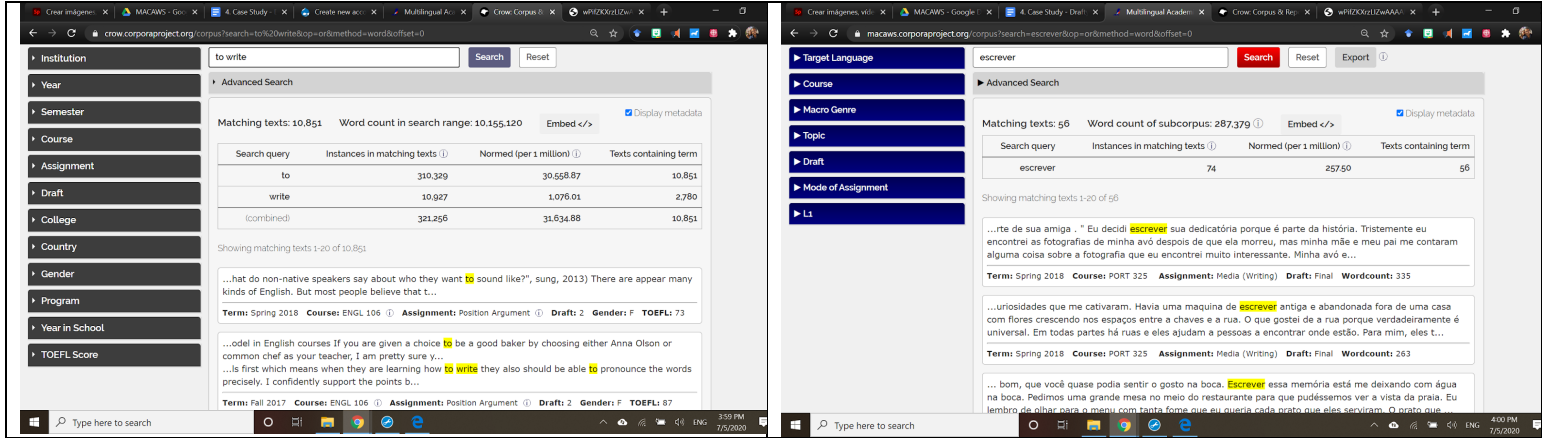*Figure 1: Screenshot of the main corpus search interface*

Figure 2: Individual page view of a pedagogical material

Case study: Adapting the software for the MACAWS corpus

MACAWS is a corpus and repository for World Languages. Though their datasets differ – most notably, MACAWS includes Russian and Portuguese language texts – Crow's search interface fit the needs of the MACAWS corpus.

In terms of data and metadata, they differ in:

Corpus & Repository of Writing https://crow.corporaproject.org/	Multilingual Academic Corpus of Assignments: Writing and Speech https://macaws.corporaproject.org
Focused on Second Language Writing, particularly English as Second Language.	Focused on Second Language Writing and Speech, particularly World Languages such as Portuguese,and Russian so far.
Repository availability: Yes! Tabs: for instructors, students, researchers, and developers available.	Repository availability: Soon! Tabs: not available.
Access: Sign in required Request access here.	Access: Sign in required Request access here.
Faceted metadata search Institution Year Semester Course Assignment: *Categorized by writing genre* Draft College Country Gender Program Year in School TOEFL Scores On Crow you can do: Keyword/phrase search Viewing individual texts Embedding search results (iDDL) Exporting results	Faceted metadata search Target language Course Macro genre Topic Draft Mode of the assignment: *Speaking or Writing* L1 On MACAWS you can do: Keyword/phrase search Viewing individual texts Embedding search results (iDDL) Exporting results