GRADUATE CERTIFICATE IN COMPUTATIONAL LINGUISTICS

The Postbaccalaureate Certificate in Computational Linguistics is designed to provide academic training in the study of computational approaches to language analysis. The curriculum assumes no prior linguistic or programming knowledge and introduces students to a variety of computational methods and their theoretical underpinnings including: writing programs in Python to process raw texts (tokenization), discovering statistical patterns in linguistic data (frequency distribution), performing part-of-speech tagging, text segmentation, and classification (context-free grammars, dependency grammars), extracting meaning from texts, and applying various machine learning methods to data mining. Computational tools are supported by a strong foundation in linguistic theory including natural language syntax, semantics, and metaphor identification.

Learning Outcomes

1. Identify grammatical categories and basic principles phonological and/ or syntactic grammar.
2. Write programs in a programming language, e.g. Python, and to process raw texts.
3. Discover statistical patterns in linguistic data, identify frequency distributions, and perform tokenization.
4. Perform part-of-speech tagging, text segmentation, and classification.
5. Build dependency grammar and extract meaning from texts.
6. Apply various machine learning methods to data mining.

Graduate Certificate in Computational Linguistics — 15 units

Core (12 units)

Code	Title	Units
ENG 620	Computational Analysis of Language Data for Artificial Intelligence (AI)	3
ENG 680	Language Models and Artificial Intelligence (AI) Applications	3
ENG 719	Seminar: Contemporary Semantic Theory	3
ENG 707	Topics in Language Analysis	3

Elective (3 units)

Select one:

Code	Title	Units
ENG 821	Syntax	3
ENG 737		

Note: A non-ENG prefixed course is allowed with prior approval from an advisor.