

We are pleased to announce that

George Mikros

has accepted the invitation of the Institute of the Czech National Corpus and is coming to Prague to give a lecture at the Faculty of Arts.



Stylometric analysis and machine learning methods

Thursday, May 16th 2019, 3:50 p.m.

Faculty of Arts (main building), 1st floor, room 18, Jan Palach Square 2

Abstract

In this lecture we will present various key developments in the Authorship Attribution (AA) methodology. More specifically we will examine a very successful and widely used stylometric feature in AA studies, i.e. the n-gram. We will investigate n-grams in character and word levels and explore their quantitative properties. We will also discuss the various methods of tokenization existed for these kinds of units and some reference to previous studies. In addition we will present an overview of the most effective machine learning algorithms used in AA (Support Vector Machines and Random Forests). This will be a non-technical presentation and the focus will be on the concepts underlying the specific algorithms.

