**The Department of English
in collaboration with
The Center for Arts and Humanities, the Faculty of Arts and Sciences,
and AUB Libraries
would like to invite you to a Workshop on
Textual Analysis Using Stylometry
(24-25 April, 2018)**

Building on the momentum created by last year's Digital Humanities Institute (DHI-B 2017) and the expansion, therewith, of the Digital Humanities community of practice within AUB and beyond, this workshop on Textual Analysis Using Stylometry will help sustain all previous iterations and efforts in the field of Digital Humanities that have come out of the Department of English in collaboration with the Center for Arts and Humanities, the Faculty of Arts and Sciences, and the University Libraries. Technologists, librarians, graduate and undergraduate students and faculty across campus will also be introduced to an important tool in the field of digital textual analysis.

About Stylometry:
Stylometry, or the study of measurable features of (literary) style, such as sentence length, vocabulary richness and various frequencies (of words, word lengths, word forms, etc.), has been around at least since the middle of the 19th century, and has found numerous applications in authorship attribution research. These applications are based on the belief that there exist such conscious or unconscious elements of personal style that can help detect the true author of an anonymous text. But even more interesting research questions arise beyond bare authorship attribution: patterns of stylometric similarity and difference also provide new insights into relationships between different books by the same author; between books by different authors; between authors differing in terms of chronology or gender; between translations of the same author or group of authors; helping, in turn, to find new ways of looking at works that seem to have been studied from all possible perspectives.

Workshop Outcomes:
The workshop will be an opportunity for participants to apply a few stylometric methods (including clustering, principal components analysis, and so on) to a collection of raw text files; they will learn how to interpret the results of a stylometric experiment; last but not least, they will be introduced to the concept of empirical inference, which, among other things, involves the notion of the reproducibility of experiments.

Workshop Leader:
The workshop will be led by Dr. Maciej Eder, an Associate Professor at the Institute of Polish Students at the Pedagogical University of Krakow, Poland, and at the Institute of Polish Language at the Polish Academy of Sciences. He is a leading expert on computational stylometry, the co-author (together with Jan Rybicki) of the stylometry packages for the R programming language, and has offered a number of workshops in the field.

Registration:
To attend this workshop, register [here](#)
Registration is free.

Schedule:
Day 1: 24 April:
2-5pm Workshop (Jafet E-Classroom, AUB)
        Round table presentations on Stylometry (20 min)
        Presentations on Integrating Stylometry in Research and Teaching (20 min)
        Hands-on session (1): (2 hrs and 20 min with short coffee break)
                installing R on laptops from the internet
                Cds/flash drives with: easy short instructions, relevant scripts and a number of 'clean' texts collections
                hands-on analysis to produce as many different results as possible
                analysis of visualizations and results.

Day 2: 25th April:
10-1pm Hands-on Session (2) with your own texts (Jafet E-Classroom, AUB)
3-5:30pm Hands-on Session (3) (optional) Fisk 204A Lab

Workshop Materials:
A corpus (a group of plain .txt files - of any author or authors you would like to study) is already available. However, participants are encouraged to bring their own. If you are bringing your own corpus, make sure that the plain text files are saved in Unicode (UTF-8), which is rather immaterial for English, but crucial for Arabic, Cyrillic etc. Each text should be saved in a dedicated file. It is convenient to name the files so that they contain some metadata, preferably separated by an underscore "_",
      E.g.: Conrad_Nostromo.txt
          Austen_Pride.txt

No laptops are required. If you do want to bring your laptop, please indicate that on the registration form and follow the installation instructions below.

Installation Instructions:
Before the workshop, participants are advised to have R installed, with the right to install packages from CRAN. If this is not possible, then launch R session and type: install.packages("stylo")
Additionally, you would need Java and Gephi.