

Stylometryczna analiza tekstu literackiego w półtorej godziny

Saturday, 20 January 2024 09:00 (1h 30m)

Wbrew pozorom to nie szyk zdania czy “wielkie”, znaczące słowa najłatwiej zdradzają, kto napisał dany tekst. Jeżeli mamy teksty kilku różnych autorów, najłatwiej pogrupować je według autorstwa licząc proporcje słów mało “znaczących”, takich jak zaimki, przymyki czy czasowniki modalne. zresztą nie tylko według autorstwa: ta sama metoda często wskazuje chronologiczną kolejność w twórczości danego pisarza. Jednym z narzędzi takich analiz jest pakiet stylo; półtorej godziny warsztatu wystarczy, by każdy z uczestników przeprowadził swoje pierwsze badanie stylometryczne.

Słowa kluczowe

stylometria, atrybucja autorstwa, analiza skupień, pakiet stylo

Nazwa projektu

Projekt Flagowy “Digital Humanities Lab”

Wymagania sprzętowe i inne

1 komputer na uczestnika

Zainstalowany R (<https://cran.r-project.org/bin/windows/base/R-4.3.2-win.exe>)

W R zainstalowany pakiet stylo

Dostępny dla użytkownika pakiet materiałów (w załączniku)

Primary author: Prof. RYBICKI, Jan (JCHC, PF Digital Humanities Lab)

Presenter: Prof. RYBICKI, Jan (JCHC, PF Digital Humanities Lab)

Session Classification: Warsztaty 2