
JGAAP - Estilometria

Autor:

Data de publicació: 02-04-2020

Welcome to the home page of Signature, a program designed to facilitate "stylometric" analysis and comparison of texts, with a particular emphasis on author identification. The collage below on the right illustrates the sorts of task for which Signature can be used: comparing the styles of Jane Austen and other novelists; examining the "authorial signature" of the plays written by (or controversially attributed to) Shakespeare; establishing the provenance of ancient manuscripts such as the shared books of Aristotle's Ethics; identifying the author of the unattributed Federalist Papers; and investigating the relationships between Biblical scriptures (e.g. Did "Luke" write Acts? Did Paul write Hebrews?).

GAAP

Software de código abierto que permite a los no expertos utilizar algunos de los últimos métodos de aprendizaje automático en sus problemas de clasificación de texto. Por favor visite JGAAP en GitHub.

Ventajas de JGAAP

Proporciona una amplia personalización analítica, como Canoniziers (normaliza textos), Culling (lo que se elimina de los datos), Analytical Events (características tales como n-grams, longitud de palabra, etc.) y métodos de análisis (Burrow's Delta, Chi-Squared, Etc).

Puede procesar múltiples textos y realizar diferentes tipos de análisis a la vez.

Funciona fácilmente con Java, no requiere software adicional ni conocimientos de codificación.

La interfaz gráfica de usuario proporciona indicaciones y orientación relacionadas con diferentes opciones estadísticas, como el sacrificio, los métodos de análisis y las opciones de frecuencia.

Desventajas de JGAAP

No genera visualizaciones de datos; Sólo genera puntuaciones estadísticas en bruto, que luego podmos procesar, por ejemplo con RStudio.

Aunque la Guía del usuario es muy amplia, no ha sido actualizado desde la versión 5.1 (2013)

Dato curioso: Juola y JGAAP son más conocidos a partir del descubrimiento de JK Rowling como autor de El canto del cuco .

ACCEDER A LA HERRAMIENTA