



## **Workflow de Técnicas de Estilometria para Atribuição Automática de Autoria**

*Diógenes D'Avila Goldoni*

*Unilasalle*

*Patrícia Kayser Vargas Mangan (Orientadora)*

*Unilasalle*

### **Tipo de Trabalho**

Pôster

### **Tema**

Engenharias e Computação

### **Palavras chaves**

Plágio, Atribuição de Autoria, Estilometria

### **OBJETIVO**

Pautado pela indubitável importância e crescente necessidade da correta atribuição de autoria de documentos textuais, em especial no meio acadêmico, este trabalho tem por objetivo a criação de um workflow para a utilização de técnicas de estilometria para a detecção automática de artigos cuja autoria é disputada por um ou mais possíveis autores.

### **MATERIAL**

Para alcançar o objetivo almejado houve a necessidade de: (1) avaliar as ferramentas e algoritmos já existentes, os quais possam ser aplicados à corpus em português, (2) determinar, empiricamente, quais as técnicas de estilometria geram maior assertividade em corpus da língua portuguesa, e (3) a implementação do workflow proposto, o qual foi utilizado de forma supervisionada a fim de obter a atribuição de autoria correta em, pelo menos, 90% dos casos.

### **METODOLOGIA**

A sistemática revisão bibliográfica de livros e artigos à cerca das técnicas de pré-processamento textual e estilometria serviram de norte para a fundamentação teórica do trabalho e para a compreensão do que já foi pesquisado e desenvolvido na área, além de prover o conhecimento necessário para a subsequente implementação prática proposta. Para a implementação do workflow, foi efetuado um estudo de caso para demonstrar a utilização das técnicas pesquisadas através da sua aplicação em



textos cujos autores já são conhecidos, de forma a verificar sua viabilidade de uso em corpus na língua portuguesa e seu grau de assertividade.

### **RESULTADOS**

O resultado obtido através da aplicação do workflow em 9 artigos científicos, de três diferentes autoras da área da educação, a fim de descobrir a real autoria de um décimo artigo em disputa fictícia foi de 100% de acurácia, predizendo e atribuindo corretamente a autoria do artigo em questão.

### **CONCLUSÕES**

Foi demonstrado que as técnicas de estilometria e pré-processamento utilizadas na criação do workflow são de grande valia para atribuição automática de autoria e o combate ao plágio acadêmico. Como sugestão para trabalhos futuros, sugere-se a exploração da assertividade do workflow em artigos com múltiplos autores.