

INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
UNIVERSIDADE DE SÃO PAULO

Proposta para Trabalho de Formatura
Supervisionado

Aluno: Lucas Ikeda França

Orientador: Prof. Roberto Hirata Jr.

Tema

Identificador de plágio para o Adessowiki.

Resumo

Plágio é uma questão presente no mundo artístico, tecnológico, e acadêmico. A apropriação indevida de trabalhos e ideias de terceiros traz problemas para diversos segmentos da sociedade, e a crescente disponibilidade de informação nos últimos anos facilita cada vez mais essa prática.

O Adessowiki é um ambiente de aprendizado colaborativo online, e permite o compartilhamento do conhecimento entre seus usuários. É possível escrever textos científicos e código fonte nas linguagens Python e C++. Porém, apesar da possibilidade de um usuário acessar o código de outros, devemos encorajar que ele compreenda a solução adotada pelo colega para então desenvolver a sua própria, e não simplesmente copiar o trabalho alheio e apresentá-lo como original. Um mecanismo que identifique tal situação pode diminuir sua incidência.

Há diversas técnicas para detecção de plágio, com diferentes graus de complexidade, que dependem do tipo de conteúdo a ser analisado. Para código fonte, podemos citar o uso de tokens e árvores de análise sintática. Já para textos, uma solução convencional é a métrica, isto é, mede-se um tipo de distância entre dois textos. O valor obtido indica o quão similares são as entradas.

O software será desenvolvido em Python, devido à sua flexibilidade e quantidade de bibliotecas existentes para diversas finalidades, como autenticação de sessão (`urllib`), análise sintática (`pyparsing`), etc.

Objetivo

O objetivo principal deste trabalho é criar uma ferramenta que auxilie na detecção de plágio no Adessowiki, a fim de inibir esta prática. Tanto o plágio de texto quanto de código fonte fazem parte do escopo deste projeto.

Atividades já realizadas

- levantamento de referências bibliográficas sobre identificação de plágio
- estudo de bibliotecas e ferramentas auxiliares
- implementação de funções básicas

Estrutura esperada da monografia

- Parte técnica
 - introdução
 - conceitos e tecnologias estudadas
 - atividades realizadas
 - resultados e produtos obtidos
 - conclusões
 - bibliografia
- Parte subjetiva
 - desafios e frustrações
 - disciplinas relevantes para o trabalho
 - próximos passos