

Genealogia Acadêmica Lattes

Edson K. Miyahara, Jesus P. Mena-Chalco, Roberto M. Cesar-Jr
Instituto de Matemática e Estatística, USP, SP

1. Objetivos

O trabalho está inserido no contexto de prospecção de dados para análise de genealogia acadêmica. Esta informação é de importante interesse para análise de inter-relação de orientações. Neste trabalho criamos uma ferramenta para geração automática da árvore genealógica para cientistas e/ou acadêmicos, cadastrados na Plataforma Lattes, através de suas relações de orientação e supervisão concluídas.

2. Metodologia

O trabalho está sendo desenvolvido em Python no ambiente Gnu/Linux. A ferramenta obtém os currículos vitae (CVs) da Plataforma Lattes a partir do ID Lattes fornecido e guarda-os em um *cache*, então processa e armazena-os em uma estrutura de dados que permite gerenciar e acessar rapidamente as informações dos CVs. Também foi criado um algoritmo para a identificação de nomes similares evitando assim a ambiguidade entre nome de uma pessoa escrita de formas diferentes. São gerados diferentes grafos em formatos que permitam a posterior visualização e análise por ferramentas externas/complementares.

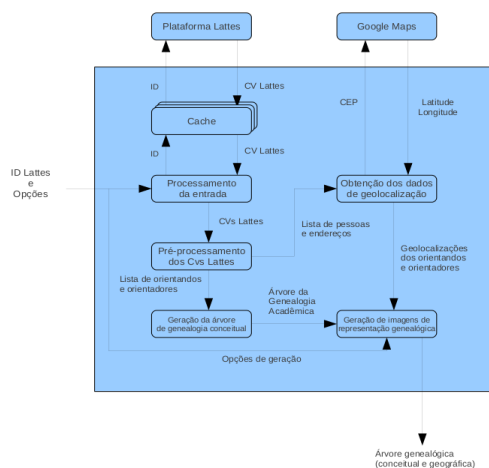


Figura 1: Fluxograma com os processos envolvidos.

3. Resultados

A ferramenta foi testada satisfatoriamente considerando mais de 100 CVs Lattes. A seguinte figura foi gerada automaticamente por nossa proposta:

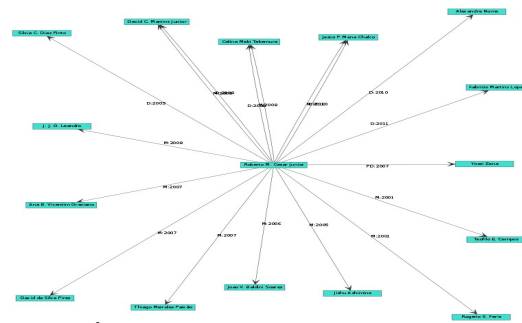


Figura 2: Árvore genealógica acadêmica do Prof. Dr. Roberto M. Cesar-Jr.

4. Conclusões

Nosso trabalho, ainda em andamento, considera diversos procedimentos para a geração automática da árvore genealógica de ascendência ou descendência de orientação acadêmica. Atualmente estamos trabalhando na: (i) geração da árvore em um mapa geográfico usando a API do GoogleMaps, e (ii) exportação de informações em arquivos de dados que permitam uma visualização e análise por ferramentas externas. É importante frisar que esta proposta é nova no Brasil e está sendo desenvolvido no IME-USP no âmbito de um TCC.

Para mais informações consulte:

<http://www.linux.ime.usp.br/~edsonkm/mac499>

5. Referências Bibliográficas

MENA-CHALCO, J. P.; CESAR-JR, R. M. ScriptLattes: An open-source knowledge extraction system from the lattes platform. *Journal of the Brazilian Computer Society*, v. 15, n. 4, p. 31-39, 2009.

Xu, G.; Zhang, Y.; Li, L. *Web Mining and Social Networking: Techniques and Applications*, Springer, 2010.