

# MAC499 – Trabalho de Formatura Supervisionado

Aluno: Edson Kiyohiro Miyahara  
Orientador: Roberto Marcondes Cesar Junior  
Co-orientador: Jesús P. Mena-Chalco

## Proposta

### **Tema**

Genealogia Acadêmica Lattes

### **Resumo**

Genealogia Acadêmica é uma tentativa de organizar em uma árvore genealógica os cientista e acadêmicos de acordo com suas relações de supervisão de relacionamentos.

Plataforma Lattes é uma base de dados de currículos de pesquisadores e instituições das áreas de ciências e tecnologias.

Web Mining é uma metodologia de recuperação de informação, que se utiliza de ferramentas de mineração de dados para extrair informações tanto do conteúdo das páginas e de sua estrutura de relacionamento, quanto dos registros de navegação dos usuários.

ScriptLattes<sup>1</sup> é uma ferramenta de extração e visualização de conhecimento a partir de currículos da plataforma Lattes.

Este projeto é uma tentativa de geração automática de árvores genealógicas para cientistas/acadêmicos, cadastrados na Plataforma Lattes, através de suas relações de orientação ou supervisão concluída. Dado apenas um único ID Lattes é gerada automaticamente a ascendência e descendência de orientação acadêmica para o pesquisador. Caso o ID Lattes do orientando seja identificado no CV Lattes do orientador, então o nó é expandido por mais um nível. A quantidade de níveis é limitado por um valor informado pelo usuário.

---

<sup>1</sup> Disponível em <http://scriptlattes.sourceforge.net/>

## Objetivos

Os principais objetivos desse trabalho são resolver os seguintes desafios:

- Armazenamento dinâmico dos CVs Lattes, com o intuito de reduzir o número de acesso à rede.
- Tratamento de nomes similares para tentar decidir se nomes escritos de formas diferentes pertencem a mesma pessoa, por exemplo, E. K. Miyahara e Edson Kiyohiro Miyahara.
- Gerar a árvore de genealogia acadêmica descendente.
- Gerar a árvore de genealogia acadêmica ascendente.
- Visualização da árvore.
- Plotar os nós da árvore no mapa geográfico

## Atividades já realizadas

- Estudo da linguagem de programação Python.
- Implementação do sistema de cache.
- Implementação de um algoritmo para verificar a similaridade entre nomes de pessoas escritas de forma diferente.

## Cronograma

	Março	Abril	Maió	Junho	Julho	Agosto	Setembro	Outubro	Novembro
Estudo/pesquisa	X	X	X	X	X	X	X		
Código		X	X	X	X	X	X	X	
Monografia					X	X	X	X	X
Pôster									X

## Estrutura esperada da monografia

- Introdução
- Conceitos e tecnologias
- Métodos de geração automática de árvores de genealogia
- Principais desafios enfrentados
- Resultados
- Conclusão
- Parte subjetiva
- Referências

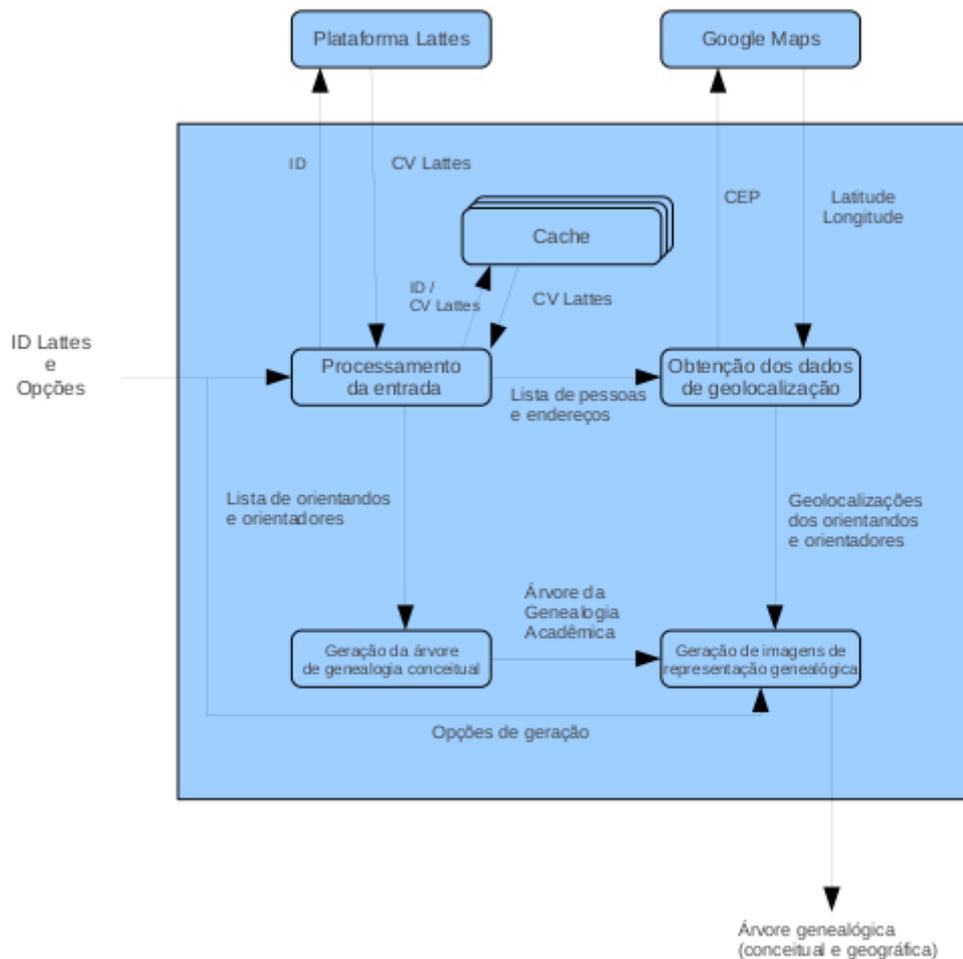


Figura: Fluxograma com os processos envolvidos.

## Referências

- [1] Genealogia Acadêmica, <[http://pt.wikipedia.org/wiki/Genealogia\\_acadêmica](http://pt.wikipedia.org/wiki/Genealogia_acadêmica)>
- [2] Plataforma Lattes, <[http://pt.wikipedia.org/wiki/Plataforma\\_Lattes](http://pt.wikipedia.org/wiki/Plataforma_Lattes)>
- [3] Web Mining, <[http://pt.wikipedia.org/wiki/Mineração\\_da\\_Web](http://pt.wikipedia.org/wiki/Mineração_da_Web)>
- [4] ScriptLattes, <<http://pt.wikipedia.org/wiki/Scriptlattes>>