

# Desenho e implementação de um sistema de análise de dados de produtividade científica

**Aluno:** Paulo Henrique Silva Araujo

**Orientador:** Marcelo da Silva Reis

Centro de Toxinas, Imuno-resposta e Sinalização Celular (CeTICS)

Laboratório de Ciclo Celular (LECC), Instituto Butantan, 13 de maio de 2020

## Resumo

A produtividade científica é avaliação da contribuição do produtor de conteúdo científico e acadêmico, ao universo do conhecimento. O Instituto Butantan agrega os marcadores individuais de seus pesquisadores em um relatório anual, preenchido digitalmente, na plataforma Scientimeter, e persistido em banco de dados. Nessa proposta, implementaremos um meio para a análise dos dados desses relatórios. Será a continuação da plataforma, desenvolvendo uma interface visual capaz de coletar e sumarizar, em gráficos e tabelas, métricas relevantes, onde os pesquisadores poderão acompanhar seus desempenhos.

# Sumário

1	Introdução	3
2	Objetivos	4
3	Metodologia	4
4	Plano de trabalho e cronograma de execução	6
	Referências	7

# 1 Introdução

O estudo científico gera produções em revistas especializadas, livros e orientações acadêmicas. Esses itens possuem avaliações qualitativas realizadas por pares ou instituições. Combinações quantitativas desses elementos podem definir índices para qualificar a produtividade científica de um pesquisador. O Instituto Butantan possui um relatório anual, a ser preenchido individualmente pelos seus pesquisadores. O Centro de Desenvolvimento Científico (CDC), do IB, é o responsável pela coleta, aprovação e encaminhamento à Secretaria da Saúde do Estado de São Paulo.

Ele era produzido em planilha Excel e continha dados de: orientações; publicações científicas e livros; aulas ministradas e participações em congressos. As planilhas digitais necessitam estar fisicamente copiadas em cada computador do usuário e depois precisam ser encaminhadas uma a uma, ao responsável. Esse precisará compilar em apenas um documento a ser enviado à Secretaria de Saúde. Além disso, existe a possibilidade de preenchimento livre, o que compromete a integridade dos dados. Por fim, não é uma solução escalável e existem ferramentas computacionais adequadas para essa finalidade. Visando ampliar a regularidade e a padronização do processo, pesquisadores do Butantan desenvolveram um sistema, chamado Scientiometer [1], para substituir a forma anterior.

O Scientiometer é um sistema web, onde os pesquisadores do Butantan inserem os dados de produções científicas e o diretor de divisão realiza a aprovação, para posterior submissão à apreciação do CDC. A utilização dele proporcionou a normalização dos dados, regularização de criação, preenchimento, envio e aprovação, ademais, propiciou armazenamento em banco de dados consistente, seguro e de simples recuperação. Deixado como trabalho futuro é o desenvolvimento de um módulo de análise e geração de interface para os

usuários, com relatórios e métricas, baseados nos dados recolhidos.

## 2 Objetivos

O objetivo é a disponibilização de um sistema para acompanhamento de resultados científicos, aos pesquisadores do Butantan, por meio da implementação de um módulo de análise de dados e produção de dashboards.

Será o cumprimento da proposição primeira do Scientiometer, de modo a computar métricas realizando consultas no banco de dados existente e apresentá-las por meio de gráficos, criando novas interfaces de usuário, na página web já utilizada.

## 3 Metodologia

O projeto de back-end, em sua comunicação com o banco de dados, deverá ser alterado para permitir as agregações necessárias, ou seja, recuperar os registros baseados nos campos das tabelas.

Além disso, as APIs, que são as interfaces com o projeto front-end, permitirão a inclusão desses campos na URL, no formato *chave=valor*.

Exemplo `https://scientiometer.com/accounts?name=Lab`

Nesse exemplo, na tabela *Account*, o campo *name* deverá conter o valor *Lab*.

O projeto front-end será alterado para efetuar requisições web, para o back-end, de forma a construir as URLs com os filtros necessários para o preenchimento dos gráficos.

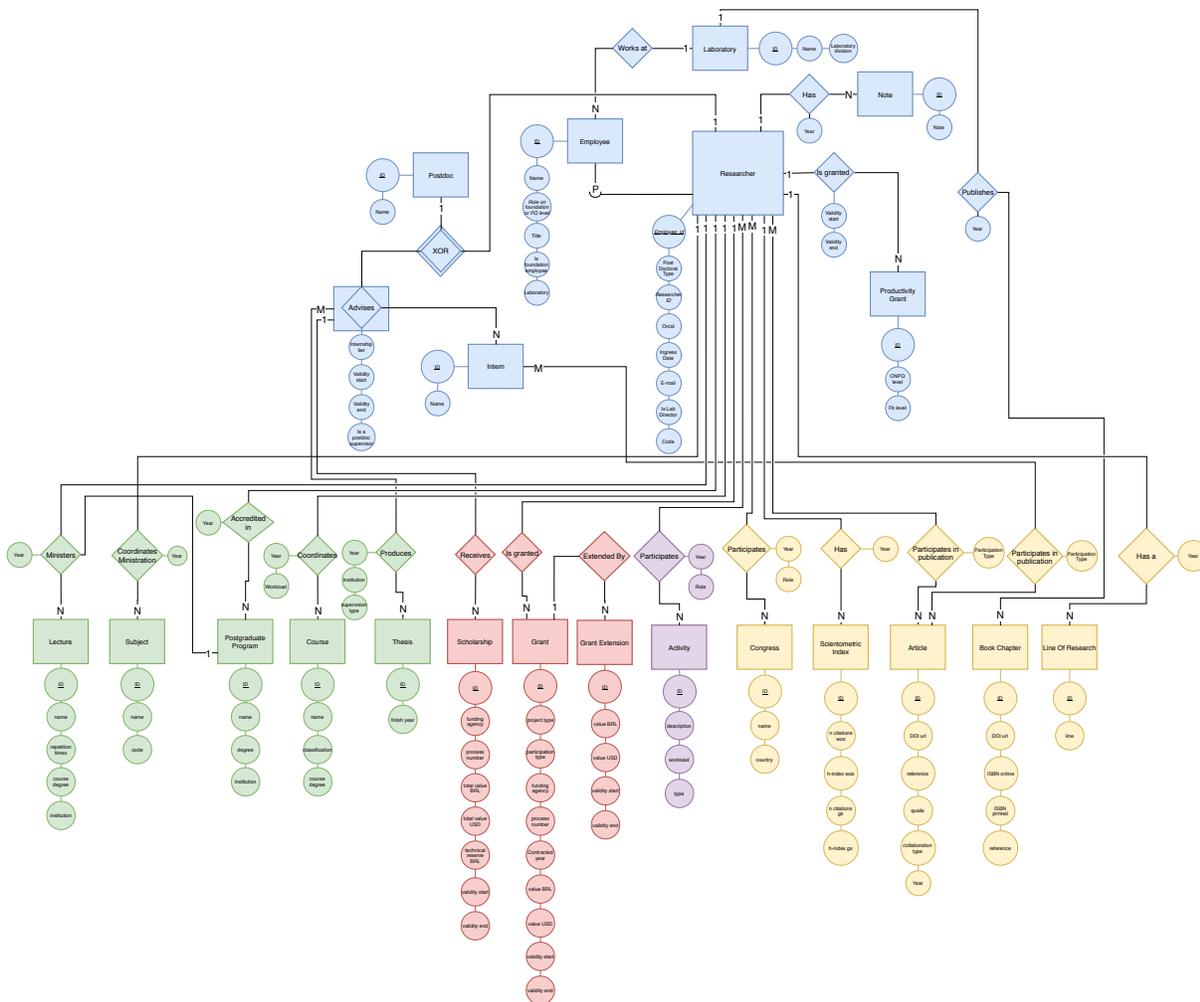


Figura 1: Modelo Entidade-relacionamento completo do sistema.[1]

A implementação visual será com Vue-chartjs [2], uma biblioteca em Vue.js, o framework de JavaScript utilizado atualmente no projeto, que fornece códigos para a criação de gráficos e tabelas.

Os pesquisadores utilizarão uma interface gráfica para selecionar quais os critérios desejados e os resultados serão exibidos dinamicamente. Esse método substituirá a forma atual, feita por consultas manuais, ao banco de dados.

Exemplo de consulta utilizada para agregar dados de fator de impacto médio:

```
SELECT division_name, laboratory_name, AVG(impact_factor) AS 'IF médio'  
FROM published_work JOIN laboratory ON (laboratory_id = laboratory.id)  
JOIN lab_division ON (lab_division_id = lab_division.id)  
GROUP BY laboratory_id  
ORDER BY division_name, laboratory_name;
```

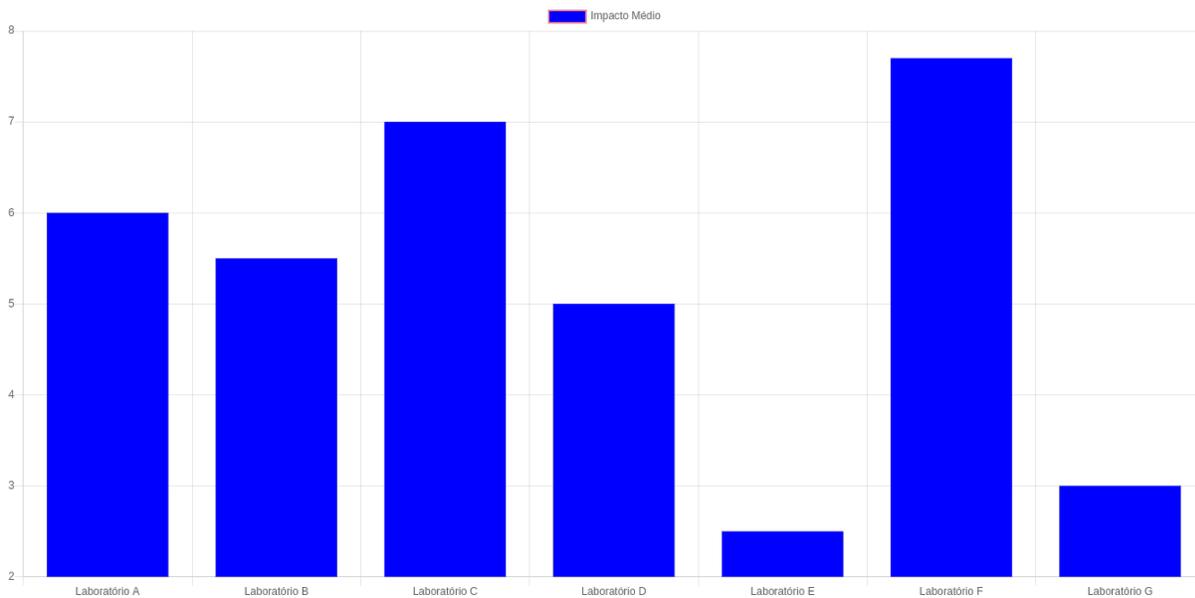


Figura 2: Exemplo de visualização desejada para impacto médio.

## 4 Plano de trabalho e cronograma de execução

Para a execução deste projeto proposto, foram listadas abaixo as principais atividades previstas. O diagrama de Gantt com o cronograma é apresentado na Tabela 1.

**Atividade 1:** Leitura inicial dos textos que servirão de base para o projeto;

**Atividade 2:** Escrita do esboço do projeto de pesquisa;

**Atividade 3:** Preparação dos dados de usuários para o ano corrente;

**Atividade 4:** Correção de bugs;

**Atividade 5:** Definição de gráficos e estatísticas desejáveis à implementação;

**Atividade 6:** Modificação do aplicativo back-end para suportar os filtros;

**Atividade 7:** Implementação de gráficos e estatísticas, no aplicativo front-end;

**Atividade 8:** Validação com usuários;

**Atividade 9:** Escrita da monografia de TCC;

**Atividade 10:** Apresentações do trabalho em pôster e, opcionalmente, seminário.

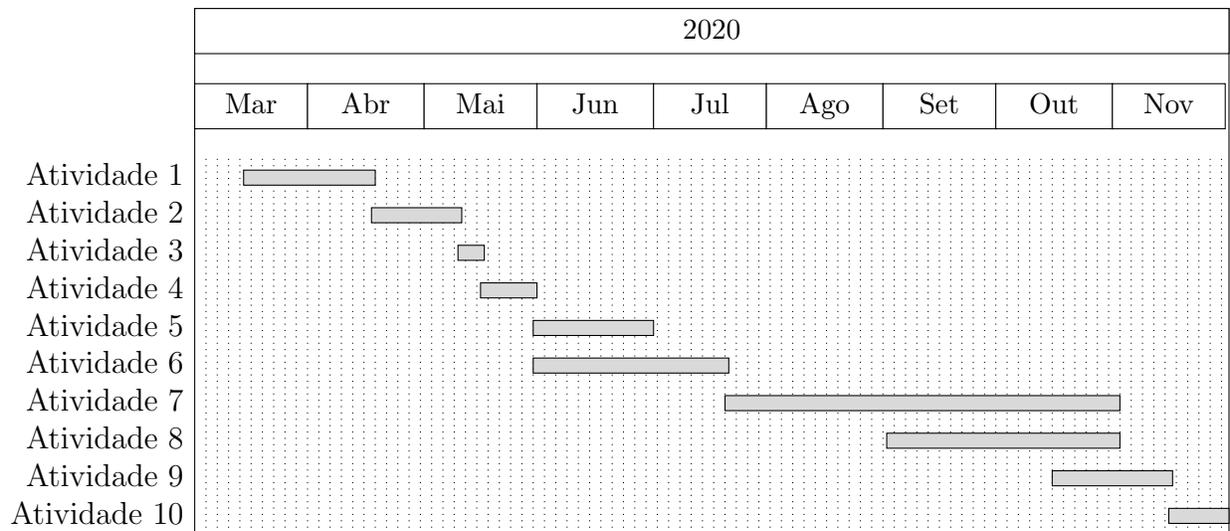


Tabela 1: Diagrama de Gantt contendo o cronograma de execução deste projeto proposto.

## Referências

- [1] Bruno B. Scholl. Desenvolvimento de um sistema de informação para coleta e análise de relatórios de produtividade científica, 2019.

[2] Vue-chartjs, easy and beautiful charts with chart.js and vue.js.