

# Um modelo computacional para construção de redes de colaboração científica em Física

Costa, C. C. S.<sup>1\*</sup>; Nascimento, J. O. do.<sup>1\*</sup>; Moret, M.A.<sup>1,2</sup>; Pereira, H.B.B.<sup>1,2</sup>

<sup>1</sup> Programa de Modelagem Computacional, Centro Universitário SENAI CIMATEC, Salvador, BA, Brasil.

<sup>2</sup> Universidade do Estado da Bahia, Salvador, Bahia, Brasil.

\*e-mail: claudinei.costa@gmail.com; jeffersonnascimento@gmail.com.

---

## Resumo

Uma forma de investigar a difusão do conhecimento na área da Física é por meio da produção intelectual dos docentes e discentes. Entender o comportamento das colaborações científicas entre estes agentes é de fundamental importância para que instituições de ensino e diferentes esferas governamentais direcionem suas ações em termos de investimento. O objetivo deste artigo é apresentar um modelo computacional alicerçado na Teoria dos Grafos, Teoria de Redes, Sistema Complexos e Engenharia de Software para a construção de redes de colaboração científica. Como uma fonte de dados em potencial para estes estudos, há a Plataforma Lattes, que é reconhecida pelo meio acadêmico para registro de produções bibliográficas. Os principais conceitos do modelo computacional são: pesquisadores responsáveis pelas criações das produções bibliográficas e a realização das orientações; os autores que participam da criação das produções bibliográficas; orientação; vértices; arestas e os grafos resultantes. O modelo computacional proposto auxilia no processo de construção de redes de colaboração científica (por meio das métricas de redes sociais e complexas) entre os diferentes níveis de pesquisadores, ajudando no processo de tomada de decisão de investimentos de instituições de ensino ou governos em áreas específicas do conhecimento em Física.

## Abstract

One way to investigate the diffusion of knowledge in the field of Physics is through the intellectual production of teachers and students. Understanding the behavior of the scientific collaborations between these agents is of fundamental importance for educational institutions and different governmental spheres to direct their actions in terms of investment. The purpose of this article is to present a computational model based on Graph Theory, Network Theory, Complex Systems and Software Engineering for the construction of scientific collaboration networks. As a potential data source for these studies, there is the Lattes Platform, which is recognized by the academic medium for recording bibliographic productions. The main concepts of the computational model are: researchers responsible for the creation of bibliographic productions and the realization of the guidelines; The authors who participate in the creation of bibliographic productions; guidance; Vertices; Edges and the resulting graphs. The proposed computational model assists in the process of building networks of scientific collaboration (through social and complex network metrics) between the different levels of researchers, helping in the decision-making process of investments of educational institutions or governments in specific areas Of knowledge in physics.

Keywords: Redes Sociais e Complexas, Engenharia de Software, UML.

---

## 1. Introdução

Conforme as atuais políticas nacionais há a necessidade de constantes verificações e avaliações

na produtividade científica nacional, por meio de discentes, docentes/pesquisadores, instituições públicas e privadas que desenvolvem atividades de pesquisa. Nesse contexto, o Currículo Lattes se

apresenta como uma referência pública na busca de informações acadêmicas de discentes e pesquisadores no Brasil. A maioria das instituições brasileiras de ensino e de pesquisa, utilizam-na como uma ferramenta de busca de informações de formação acadêmica, vínculos empregatícios, pesquisas e publicações realizadas por um determinado pesquisador.

Diante das informações disponíveis no Currículo Lattes é possível realizar estudos que podem apoiar no entendimento do comportamento da produtividade e colaboração científica dos pesquisadores e dos centros de pesquisas que os mesmos estão relacionados. A partir da observação destes comportamentos, inferências podem ser realizadas, como por exemplo: publicações em eventos e periódicos de um pesquisador, o nível de colaboração entre instituições de pesquisa e autores e se as publicações estão com uma tendência disciplinar ou interdisciplinar.

Obter dados da plataforma lattes e identificar padrões para posteriores análises é um exercício custoso pois os dados são disponibilizados de forma semiestruturada não favorecendo a análise e identificação de comportamentos. Sempre que é necessário executar análise sobre os dados mantidos nesta plataforma, se faz necessário dispensar grande esforço manual para recuperar e tratar os dados e em seguida estruturar estes dados em um formato que possibilite a realização de análises.

Este trabalho apresenta uma modelo computacional, que pode ser entendido como uma ferramenta de sistema de informação, que possibilitará as instituições de ensino:

1. Importar os dados de produções bibliográficas de seus docentes e discentes mantidos na plataforma lattes;
2. Importar dados de produções em formatos de dados semiestruturados (txt, csv, xls, pdf dentre outros);
3. Manter estes dados em base de dados local e de fácil acesso;
4. Prover métodos automatizados de tratamento dos dados;
5. Estruturar os dados de produções bibliográficas e;
6. Criar redes de colaboração científica.

## 2. Referencial teórico

A Engenharia de Software tem objetivo à aquisição de conhecimentos para a construção de produtos de software através do uso de processos, métodos e ferramentas que possibilitam melhor modelagem, desenvolvimento, operação e manutenção destes produtos.

Segundo [1], Engenharia de Software é o estabelecimento e emprego de princípios de engenharia de modo a ter, como resultado, software de maneira econômica, confiável e que funcione de forma eficiente em máquinas no ambiente de produção. [2] entende que Engenharia de Software é uma disciplina que engloba todos os aspectos relacionados com a produção de software através do uso de técnicas e ferramentas adequadas.

Dentre as técnicas e ferramentas utilizadas para projetos de software podemos destacar a Linguagem Unificada de Modelagem (UML) e a criação de protótipos.

A UML se ocupa traduzir os requisitos de um software em modelos oferecendo recursos que permitem a especificação, visualização e construção de artefatos de software que dão suporte necessário para a validação de um modelo de software. A UML fornece diagramas que apresentam os aspectos estruturais e comportamentais de um software apresentando a arquitetura necessária para o desenvolvimento da solução computacional [3].

O processo de criação de protótipos apoia a atividade de especificação de requisitos facilitando o entendimento das funções do software tanto por parte de quem demanda quanto por parte de quem desenvolve, eliminando riscos por falha de entendimento [1] e [2].

## 4. Metodologia

Segundo WAINER [4] pesquisas em ciência da computação podem ser realizadas por meio de: pesquisas analíticas, qualitativas, quantitativas e bibliográficas.

O desenvolvimento deste trabalho foi realizado com base na pesquisa quantitativa e qualitativa.

A pesquisa qualitativa foi realizada com o auxílio de entrevistas com um coordenador de um programa de pós-graduação com o objetivo de identificar o processo de análise de colaboração científica utilizado atualmente para que este sirva como balizador para construção da proposta de ferramenta. Apresentamos na tabela 1 os macros requisitos da ferramenta.

Tabela 1. Lista de macros requisitos.

Código	Descrição
RNF01	Manutenção de dados em sistema gerenciador de banco de dados relacional
RNF02	Linguagem de desenvolvimento orientada a objetos
RF01	Registro de Pesquisadores
RF02	Registro de Autores
RF03	Registro de Orientações
RF04	Registro de Produções Bibliográficas

RF05	Importação de dados de Produções Bibliográficas de Pesquisadores via currículo Lattes
RF06	Importação de dados de Orientações realizadas por Pesquisadores via currículo Lattes
RF07	Criação de Redes de Colaboração Científica.
RF08	Integração com o software especialistas para análise de redes Pajek
RF09	Integração com o software especialistas para análise de redes Gephi

A partir dos resultados das entrevistas foram construídos protótipos para apoiar a especificação dos requisitos em seguida foi realizada análise dos documentos semiestruturados (currículos lattes formato XML) disponibilizados pela plataforma Lattes com os dados de produção científica de pesquisadores, com o propósito de entender a sua estrutura e identificar quais dados disponíveis são relevantes a partir da verificação dos requisitos levantados. De posse destes dados construímos o modelo conceitual da modelo computacional proposto conforme apresentamos na figura 1.

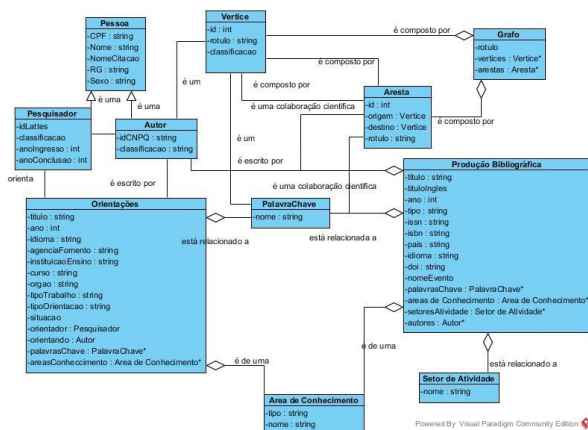


Figura 1: Modelo Conceitual. Fonte: Autores.

Após o levantamento dos macros requisitos e definição dos conceitos, foi definido o fluxo de operação da ferramenta, destacando as atividades que devem ser feitas com o uso do modelo computacional proposto e as atividades que devem ser feitas a partir da Plataforma Lattes. Apresentamos na figura 2 o fluxograma de operação.

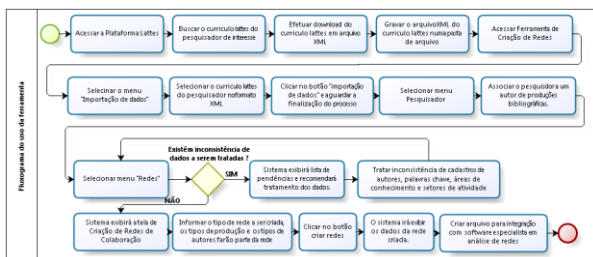


Figura 2 Fluxograma de operação. Fonte: Autores.

## 5. Resultado e discussões

Após o entendimento dos requisitos e criação de modelo conceitual, foi possível desenvolver o modelo computacional para construção de redes de colaboração científica, nomeado de Sistema de Análise de Colaboração Científica (SACC).

O SACC é um modelo computacional genérico que pode ser utilizado para construção de redes de qualquer área do conhecimento tendo sido inicialmente utilizado na sua validação para análise de um programa de pós-graduação da área de conhecimento Interdisciplinar.

O SACC permite ao pesquisador importar dados de pesquisadores suas produções bibliográficas, orientações e participações em bancas tendo como origem dos dados os currículos lattes no formato XML (de acordo a plataforma Lattes). Apresentamos na figura 3 a funcionalidade de importação de currículos lattes para o SACC.

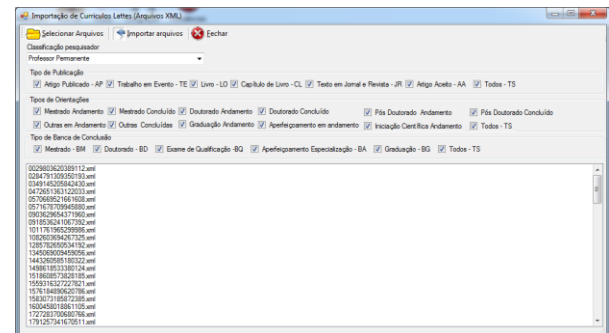


Figura 3: Importação de Currículos Lattes. Fonte: Autores.

Após a importação dos dados o pesquisador poderá criar redes uni-modais de colaboração, palavras chaves e áreas de conhecimento bem como redes bimodais de pesquisador X área de conhecimento, pesquisador X palavra chave e pesquisador X periódico tendo a possibilidade de segmentar os dados da rede por: tipo do pesquisador; período das produções e tipo das produções (artigos, capítulos de livro, livros, trabalhos em eventos e textos em jornais e revistas). Apresentamos na figura 4 a funcionalidade de criação de redes.

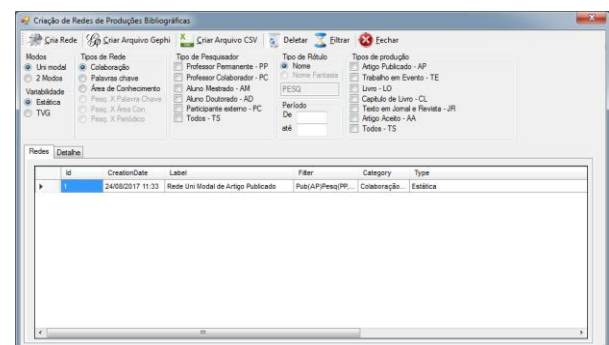


Figura 4: Criação de redes. Fonte: Autores.

O SACC, se ocupa apenas da geração, tratamento criação de redes. A tarefa de análise das redes é realizada a partir da integração com os softwares especialistas para análise de redes complexas Gephi e Pajek. Esta integração é realizada por meio de criação de arquivos em formatos lidos pelos softwares citados. Apresentamos nas figuras 5, 6 e 7 redes de coautoria, palavras chave e publicações em periódicos cuja a criação da rede foi realizada por do uso do SACC e a visualização e análise da rede foi realizada por meio do software especialista Gephi.

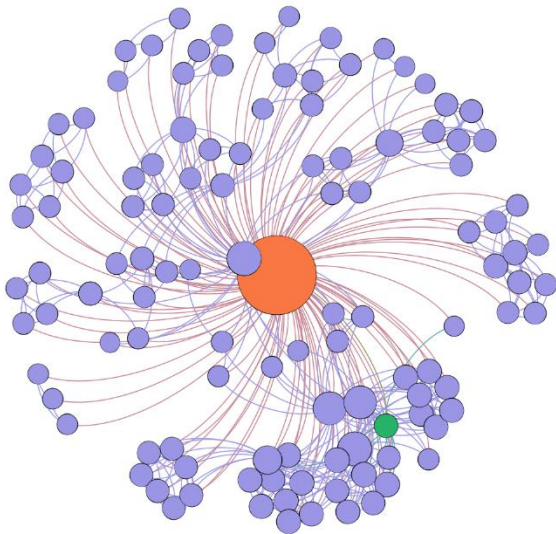


Figura 5: Rede de Coautoria. Fonte: Autores.

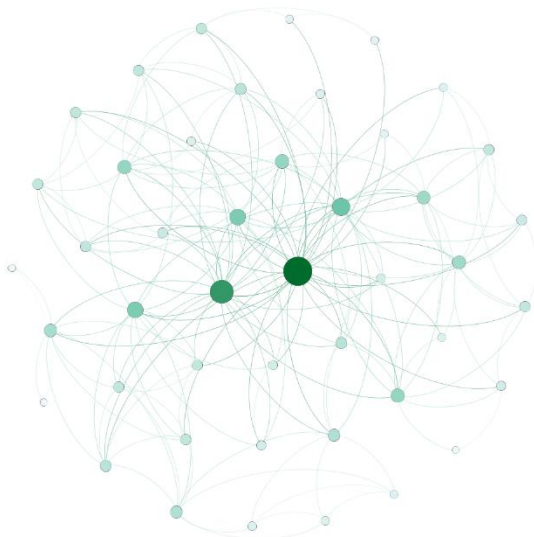


Figura 6: Rede de Palavra Chave. Fonte: Autores.

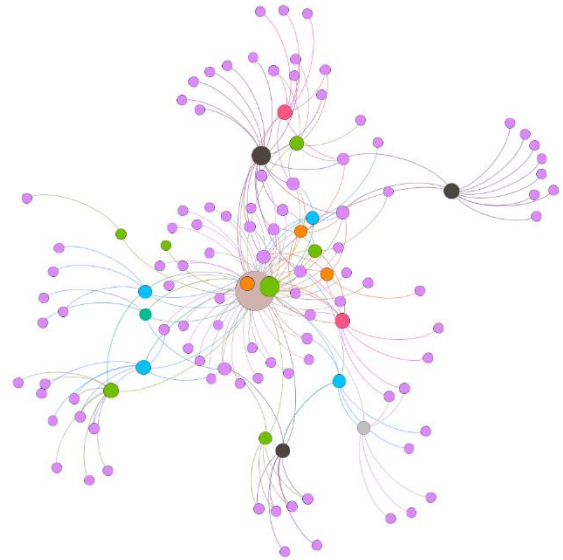


Figura 7: Rede Publicações X Periódico. Fonte: Autores.

## 6. Conclusão

Com o SACC esperamos contribuir com o processo de análise de colaboração e produções científicas entre pesquisadores de uma instituição de ensino, uma vez que facilitamos as tarefas geração de dados e criação das redes. Especialmente, consideramos a colaboração científica em programas de pós-graduação, de modo que os gestores dos programas possam direcionar as ações para estratégias que visem ampliar a colaboração científica.

## 7. Referências

- [1] Pressman, Roger S., Engenharia de Software Uma Abordagem Profissional, Tradução da Sétima Edição. São Paulo: Bookman, 2011.
- [2] Sommerville, Ian, Engenharia de Software, Tradução da Nona Edição. São Paulo: Person Education, 2011.
- [3] Rumbaugh, James; Booch, Grady; Jacobson, Ivar Uml - Guia do Usuário - Tradução da 2ª Edição. São Paulo: Elsevier, 2006.
- [4] Métodos de pesquisa quantitativa e qualitativa para a Ciência da Computação. Atualização em informática, Rio Grande do Sul. Disponível em < <http://www.pucrs.br/famat/viali/mestrado/mqp/material/extos/Pesquisa.pdf> >. Acesso em: 07 Ago. 2017.
- [5] Larman, Craig. Utilizando UML e Padrões – Tradução da 3ª Edição. São Paulo: Bookman, 2007.