

MODELAGEM CIENTÍFICA E APLICAÇÃO DE UM FRAMEWORK PARA ANÁLISE DE REDES DE COOPERAÇÃO CIENTÍFICA

SCIENTIFIC MODELING AND APPLICATION OF A FRAMEWORK FOR THE ANALYSIS OF SCIENTIFIC COOPERATION NETWORKS

José Damião de Melo^a

José Karam Filho^b

Suely Aldir Messeder^c

RESUMO

Objetivo: Este artigo tem como objetivo apresentar uma proposta de framework conceitual/operacional para a descoberta de conhecimento por meio da modelagem científica de redes de cooperação científica. **Metodologia:** Foi utilizado um caso de estudo no domínio da educação profissional e tecnológica para validar o framework proposto. O processo envolveu a extração, tratamento, carga e análise dos dados, bem como a publicação dos resultados com amplo acesso às informações, características, indicadores e métricas da rede de colaboração alvo. **Resultados:** O framework foi validado a partir de um caso de estudo, executado no domínio da educação profissional e tecnológica, mostrando-se eficaz para a extração, tratamento, carga e análise dos dados, permitindo a publicação de resultados do domínio de interesse, com amplo acesso às informações, características, indicadores e métricas da rede de colaboração alvo. **Conclusões:** A modelagem científica de redes de cooperação científica, utilizando o framework proposto, é uma estratégia promissora para a descoberta de conhecimento nesse contexto. A aplicação de técnicas de análise de redes sociais e o uso de soluções de tecnologia da informação e comunicação permitem explorar de forma eficiente as relações e os indicadores estatísticos das redes de colaboração.

Descritores: Modelagem Científica. Colaboração Científica. Redes Sociais. Educação

^a Doutor em Difusão do Conhecimento pela Universidade Federal da Bahia (UFBA). Docente da Universidade Federal da Bahia (UFBA) e do Instituto Federal de Sergipe (IFS), São Cristóvão, Brasil. E-mail: damiaomelo@gmail.com.

^b Doutor em Engenharia Mecânica pela Universidade Federal do Rio de Janeiro (UFRJ), Petrópolis, Brasil. E-mail: jkfi@Incc.br.

^c Doutora em Antropologia pela Universidad de Santiago de Compostela (USC) validado pela Universidade Federal da Bahia (UFBA). Docente da Universidade do Estado da Bahia (UNEB), Salvador, Brasil. E-mail: suelymesseder@gmail.com.

Profissional e Tecnológica.

1 INTRODUÇÃO

É uma demanda atual compreender como, onde e de que forma a cooperação e o compartilhamento de saberes acontecem na comunidade científica, como se dá o jogo dos agentes que buscam produzir ciência colaborativamente, no contexto de globalização das relações entre pesquisadores e estruturas criadoras de conhecimento, desafiando-nos a propor estratégias e ações que permitam identificar e quantificar como e de que forma ocorre essa ação colaborativa, no processo de ação dos pesquisadores.

A Análise de Redes Sociais transfere a unidade básica de análise do indivíduo, ou de um grupo de indivíduos, para um domínio mais abrangente. Introduce um componente com uma plasticidade própria do comportamento dos agentes, simbolicamente capturado e apresentado pela estrutura reticular em suas dimensões de representação. Essa abordagem leva a um dinamismo comportamental, ao mesmo tempo que implica em arranjos reticulares específicos para sua análise, pela adição do componente temporal como estruturante das relações sociais presentes na rede, fugindo ainda da exigência da permanência e admitindo capturar sua dinâmica de comportamento, permitindo “regularidades em movimento” (Fialho, 2020, p. 55).

Em termos de indicadores de colaboração científica, a produção bibliográfica em coautoria assume papel especial para a construção dessas redes, considerando que os dados necessários para aferir a totalidade sobre como, quando e de que forma ocorre a participação dos pesquisadores durante todo o tempo que durar o processo de condução de suas pesquisas, via de regra, não estão disponíveis.

As indicações de coautoria são fontes de registro histórico quanto a cooperação entre pesquisadores, permitindo inferir que houve a participação dos autores em alguma fase das atividades desenvolvidas pelos cientistas em campo, laboratórios ou grupos de pesquisa.

Mesmo sendo um indício coletado em uma etapa posterior ao processo

de pesquisa, a presença do registro bibliográfico da coautoria de produções científicas tem sido utilizada como um indicador consistente da existência de colaboração entre os pesquisadores que assinam a produção (Newman, 2001a; Mena-Chalco e Cesar Junior, 2009; De Solla Price, 1965; Aria; Cuccurullo, 2017; Dias *et al.*, 2020; Andrade *et al.*, 2017).

A modelagem de redes de cooperação científica normalmente se apoia na Teoria de Grafos, que será aqui utilizada. Um grafo é uma abstração matemática composta por dois conjuntos, o primeiro formado por pontos, chamados de vértices e o segundo conjunto formado pelas arestas, que unem pares de pontos, representando a relação existente entre os vértices (Boaventura Netto; Jurkiewicz, 2017).

Um grafo ilustra a relação existente entre os elementos presentes no sistema, permitindo representar de forma muito poderosa o problema físico de interesse. Especificamente para o caso da cooperação científica, a partir de dados de coautoria, os vértices representam os autores e as arestas representam a presença de produções bibliográficas produzidas conjuntamente, valoradas em função do quantitativo de produções em comum (Newman, 2001b; Freitas, 2010).

A capacidade dos grafos em representar abstratamente, de forma eficiente, problemas em domínios diversos foi determinante, por exemplo, para Euler resolver o problema das pontes de Königsberg, para Cayley representar a formulação química dos hidrocarbonetos, para Kirchhoff desenvolver a teoria para calcular propriedades dos circuitos elétricos e seus fenômenos característicos, para Newman analisar as relações de cooperação entre pesquisadores e para Mena-Chalco e Cesar-Junior explorarem as relações de produtividade científica dos pesquisadores brasileiros a partir dos dados da Base Lattes (Freeman, 2004; Boaventura Netto; Jurkiewicz, 2017).

A exploração dessa estratégia está presente em campos e áreas que vão desde a bibliometria (Aria; Cuccurullo, 2017), passando por ciências sociais (Higgins; Ribeiro, 2018) e educação (Leite *et al.*, 2014), com abordagens várias, apoiadas em indicadores cientométricos, com evidente destaque para medição da produção bibliográfica em coautoria e utilizando técnicas de análise de redes

sociais, abrindo uma nova frente de ação e pesquisa, onde se destacaram os trabalhos precursores De Solla Price (1965) e de Newman (2001a).

Os avanços recentes em termos de modelos e aplicações da análise de redes sociais não seriam tão evidentes sem o correspondente avanço e desenvolvimento da capacidade de armazenamento de dados e processamento computacionais, obtidas a partir da evolução da arquitetura de sistemas computacionais, da internet, da digitalização do processo de arquivamento e distribuição dos dados bibliográficos das produções científicas e da busca pela difusão do conhecimento mediado por tecnologia da informação e comunicação.

Neste artigo buscaremos contribuir com esse campo de pesquisas, apresentando uma proposta de um *framework* processual computacional de uso geral para a modelagem, geração e publicação de redes de cooperação científica, que permita a adaptação do processo de descoberta de conhecimento para múltiplos domínios, aumentando a possibilidade de representação e difusão do conhecimento científico.

Estratégias para a medição e avaliação do processo de colaboração nas diversas áreas do conhecimento possuem relevância e justificam-se como uma forma de obtenção de melhorias na ação do pesquisador, quer sejam de interesse fenômenos relativos a redes sociais de cooperação e dos modelos decorrentes, quer sejam de interesse outros domínios onde o *framework* se mostre viável, com os devidos ajustes, aprimorando a possibilidade de compreensão, análise e divulgação dos resultados das pesquisas.

Além desta seção introdutória, o artigo está organizado em três seções. A primeira apresenta o plano de fundo conceitual para a proposta, embasado na modelagem científica. A segunda seção traz a proposta de *framework* em detalhes e discute sua validação a partir da execução de um estudo de caso com aplicação a uma instituição de ensino e pesquisa. Ao final tecemos comentários e conclusões.

2 MODELAGEM CIENTÍFICA DO FRAMEWORK PARA ANÁLISE DE REDES DE COOPERAÇÃO CIENTÍFICA

A questão acerca do significado, limites e postulações que permeiam e

atravessam a própria etimologia e a compreensão dos termos “modelo” e “modelagem”, será brevemente discutida, levando em conta sua desambiguação, sua articulação no sentido e uso proposto, considerando o contexto de aplicabilidade dos termos com relação a redes de cooperação científica.

A modelagem, enquanto ação geradora de modelos, pode ser compreendida como uma atividade processual, iterativa e evolutiva para representar determinado sistema, para resolução de problemas, criando uma possibilidade de ação que era dificultada pelo sistema original que motivou construir o modelo.

Para a construção de um modelo é adequado adotar um esquema metodológico que permita ao pesquisador/modelador exercer suas habilidades em formular ou modelar um problema cientificamente, seguindo uma sistemática para articular seu arcabouço cognitivo e sua linha de raciocínio.

Os modelos são utilizados para representar uma determinada realidade ou sistema, mediante uma estrutura conceitual abstrata ou fundada em dados empíricos, com objetivo de estudar determinados fenômenos ou comportamentos que apresentam algum tipo de restrição ao seu estudo de outra forma, dada sua natureza complexa ou mesmo inacessível, por limitações de naturezas éticas, sociais ou físicas (Karam-Filho; Almeida, 2013).

Modelo é conceituado aqui no sentido do produto gerado ou desenvolvido a partir de um processo sistemático de Modelagem que, por sua vez, é entendida como a prática de substituir um sistema por outro que o represente. Essa substituição implica não somente em assumir, a priori, a necessidade de definição de uma escala de representação, mas também, em uma concepção da percepção de algo ou alguma situação que venha a ser objeto da finalidade objetiva do modelo, alvo da modelagem, item de interesse de estudo e que vem a ser o nosso fenômeno de interesse (Karam-Filho, 2019).

O modelo deverá representar a visão do modelador acerca da realidade material ou abstrata que o mesmo consegue captar, sendo dependente do repertório cognitivo e instrumental do modelador, de restrições éticas, das características observáveis do sistema físico, do domínio cognitivo, das

intenções do observador, das condições de observação e de acesso ao sistema que se deseja modelar, fazendo uso de instrumentos baseados em diversas técnicas e linguagens, sem ser limitante ou gradativo quanto às possibilidades de representação (Karam-Filho, 2019; Burnham, 2016; Galeffi, 2020; Boaventura Netto; Jurkiewicz, 2017).

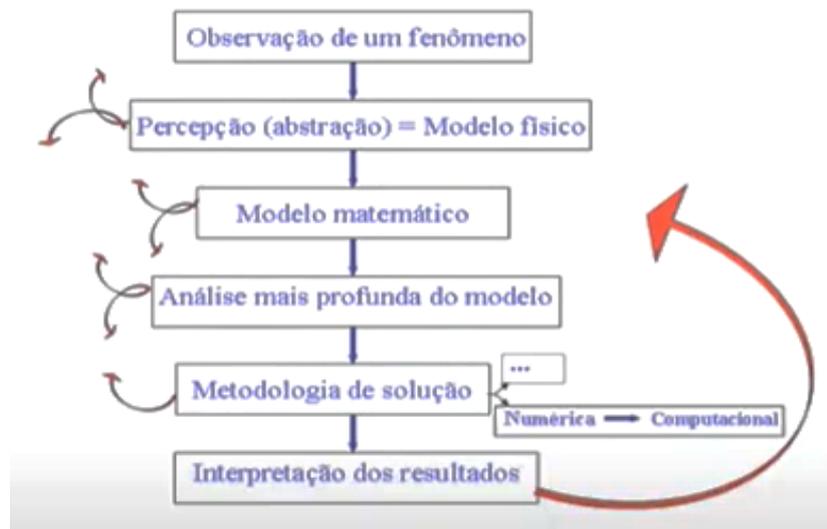
Levando em consideração a recursividade do processo e a natureza do problema abordado, a modelagem científica é o procedimento científico de representar matematicamente fenômenos ou problemas advindos das diversas áreas em formulações matemáticas adequadas, que sejam tratáveis e que forneçam ideias, respostas e informações sobre o fenômeno real observado (Karam-Filho, 2019).

Para organizar a estratégia de construção de nosso *framework*, adotamos um esquema para modelagem estruturado em seis passos ou etapas, relacionados entre si e apresentados em uma estrutura em camadas iterativas e realimentadas, cuja ação iterativa é condicionada aos resultados obtidos em cada execução parcial do ciclo de modelagem.

O modelo comporta adaptações, ampliações e ajustes, inclusive de reanálise, conforme os estágios de execução. Isto permite ao modelador ampliar seu domínio cognitivo, articulando habilidades com o conhecimento prévio e a estrutura do processo de construção do modelo, aumentando a capacidade do modelo de representar o sistema original, a partir do esquema de modelagem matemática/computacional apresentado na Figura 01.

A modelagem científica para a análise de redes de cooperação científica leva em conta conceitos e práticas relacionados com lógica e matemática, estatística, sistemas de banco de dados e mineração de dados. Estes elementos fornecem a base conceitual para soluções dessa natureza, que atuam como condicionantes para o *framework* conceitual/operacional proposto.

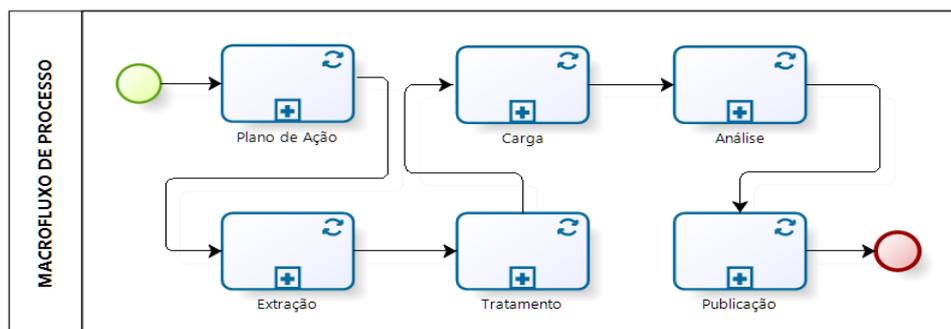
Figura 1 – Esquema para modelagem matemática/computacional



Fonte: Karam-Filho (2019).

O *framework* é formado por cinco etapas: extração, tratamento, carga, análise e publicação, que podem ser executadas em um fluxo de encadeamento de tarefas, organizadas em etapas específicas para cada ação, conjugadas em um acrônimo aqui denominado de ETCAP – Extração, Tratamento, Carga, Análise e Publicação.

Figura 02 – Macrofluxo de Processo ETCAP



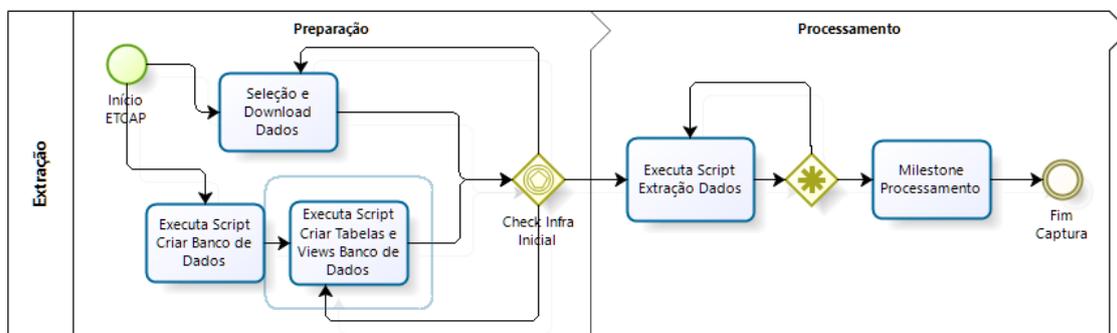
Fonte: Elaborado pelos autores.

A possibilidade de adaptação e reuso do *framework* permite a reprodução das ações em outros contextos de colaboração científica, tendo em seu percurso metodológico uma etapa inicial de ajustes e definição do domínio, seguida de

cinco etapas de ataque ao problema de pesquisa, destacados na Figura 02. Estas etapas são definidas a seguir.

Extração - Execução de tarefas relacionadas com a aquisição dos dados onde estão localizadas as informações de interesse. No caso de fontes bibliográficas, as informações são disponibilizadas em formatos não estruturados, na forma de arquivos de texto ou em notações de marcação, a exemplo do currículo da base de currículos Lattes, que é exportado individualmente em um formato com notação específica, não padronizado. A Figura 03 ilustra esta etapa.

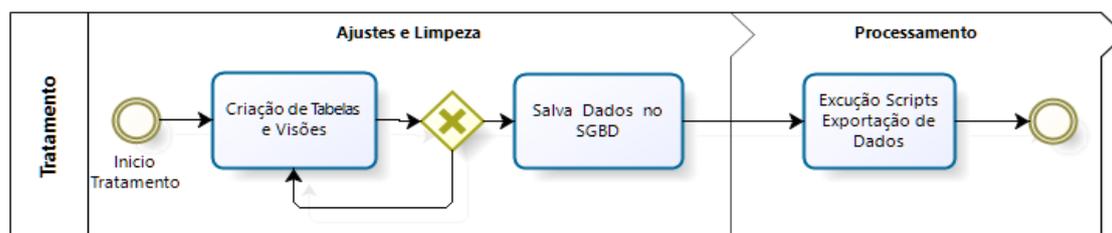
Figura 03 – Fluxo do processo de Extração



Fonte: Elaborado pelos autores.

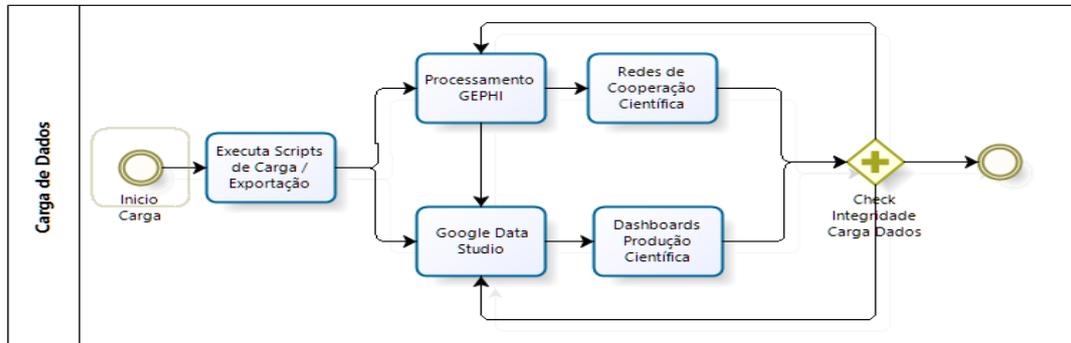
Tratamento - Neste subprocesso do framework ETCAP, as informações presentes nos documentos não estruturados são extraídas de acordo com as definições de interesse e das restrições do domínio, sendo agrupadas em pares de atributo-valor, para servir de insumo para o subprocesso de carga de dados, como ilustra a Figura 04.

Figura 04 – Fluxo do processo de Tratamento



Fonte: Elaborado pelos autores.

Figura 05 – Fluxo do processo de Carga

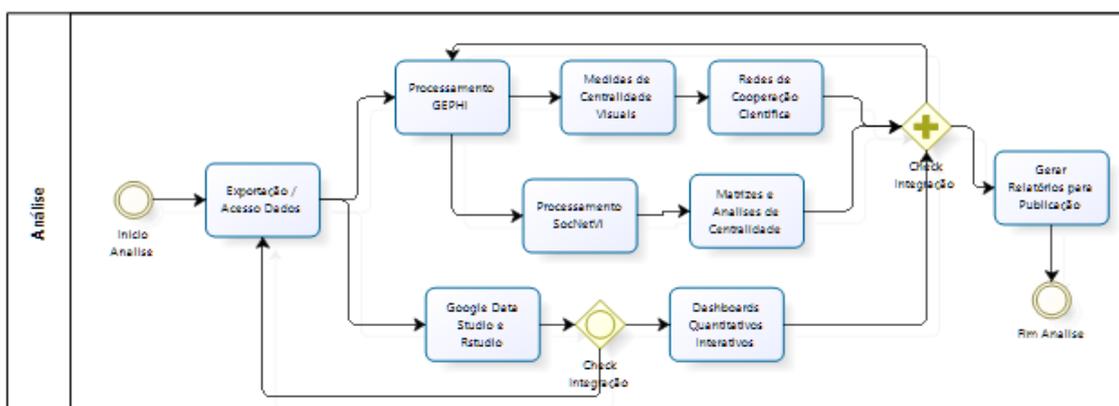


Fonte: Elaborado pelos autores.

Carga – Este conjunto de tarefas realiza o armazenamento das informações extraídas em uma nova base de dados, utilizando um sistema de gerenciamento de banco de dados. Nesta etapa de execução do framework, os dados já estarão colecionados e permitirão análises de contextualização geral da amostra, a depender do interesse e do objetivo, com a possibilidade de uso de técnicas de agregação, segmentação e classificação, típicos da estatística descritiva. A Figura 05 ilustra esta etapa.

Análise - As tarefas desta etapa estão relacionadas com a aferição acerca das informações disponibilizadas nos portais de acesso e seu grau de assertividade, considerando os objetivos e o problema para o uso do framework, ilustrado na Figura 06.

Figura 06 – Fluxo do processo de Análise



Fonte: Elaborado pelos autores.

Publicação - O subprocesso de publicação foi modelado considerando a necessidade de preservação de informações sensíveis dos agentes cooperantes, criando um nível de acesso autorizado aos dados, estruturas de computação e programas utilizados para o processamento, a partir da separação física e lógica entre os componentes desta etapa com as etapas anteriores.

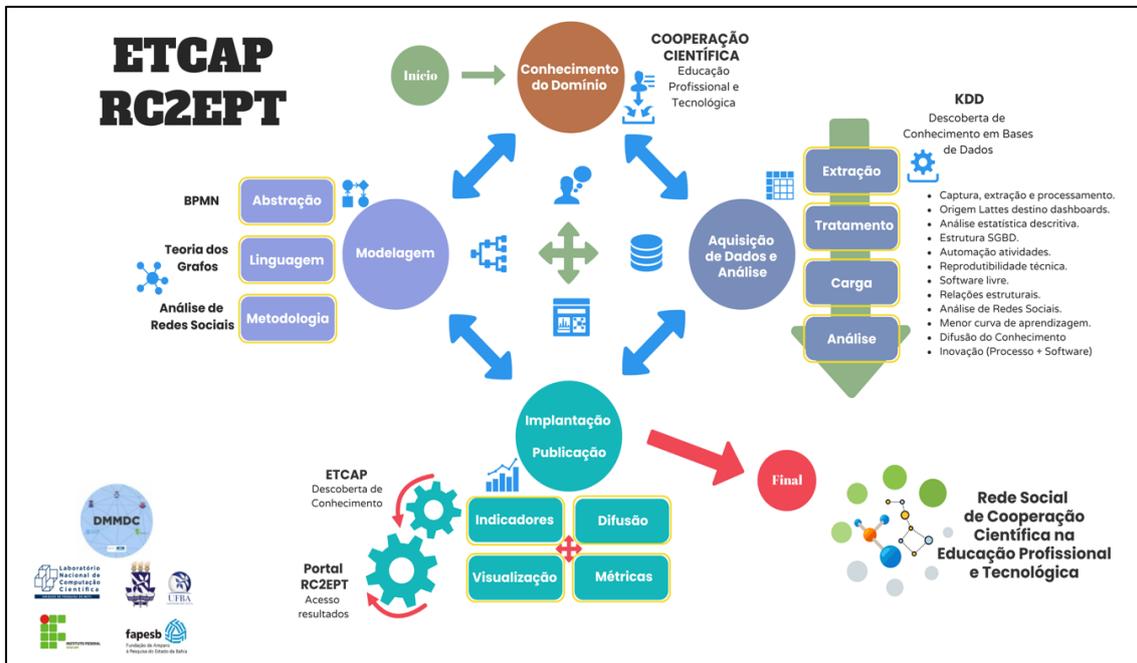
3 APLICAÇÃO DO *FRAMEWORK* ETCAP PARA NAÁLISE DE REDES DE COOPERAÇÃO CIENTÍFICA

A validação empírica do framework modelado se deu pela condução de um estudo de caso no domínio da Educação Profissional e Tecnológica, tendo como origem de dados as informações disponíveis nos currículos Lattes dos integrantes de grupos de pesquisa vinculados ao Instituto Federal de Sergipe.

Adotamos a técnica de análise de rede social completa, com a construção das redes de coautoria (Newman, 2001a), aferindo as medições de centralidade e de estrutura (Freeman, 1979; Lemieux; Ouimet, 2012), associadas a uma análise bibliométrica (Digiampietri *et al.*, 2012; Mena-Chalco; Cesar Junior, 2009), tratando as informações sobre as relações dos atores da população, sua segmentação em grupos de interesse e as relações reticulares (Fialho, 2020), analisando sua topologia e representando o panorama do domínio acerca da evolução e da cooperação entre os atores da rede (Hanneman; Riddle, 2005).

O conjunto de informações armazenadas no novo banco de dados foi formado a partir de dados extraídos do curriculum Lattes/CNPQ, tendo como referência as informações cadastradas na base de dados do diretório de grupos de pesquisa – DGP/CNPQ, considerando um recorte temporal de 2008 a 2018 (Gomes; Cavalcante; Lins, 2016).

Figura 07 – Cartografia do Framework ETCAP – RC2EPT



Fonte: Elaborado pelos autores.

Para o processamento computacional dos dados capturados desenvolvemos um conjunto de softwares e rotinas originais, desenvolvidos para uso a aplicação do framework ETCAP, voltados especificamente para a automação das tarefas de criação do banco de dados e da extração dos dados dos currículos lattes, efetuando diretamente a carga dos dados no sistema de gerenciamento de banco de dados, conforme ilustrado na Figura 07.

A nossa abordagem para atribuição de valor à relação de colaboração da rede considera que dois pesquisadores que tenham produzido um ou mais trabalhos em coautoria indicam sua participação em uma ação colaborativa e estão conectados, na medida da quantidade de trabalhos conjuntos.

Foram mantidos para a construção e análise de redes os dados extraídos dos currículos que de fato contribuíram para rede de cooperação, excluídos os nós isolados. Aplicando o critério de corte, foram excluídos da amostra 233 nós com grau menor que 1, o equivalente a 45,8% dos vértices, ou seja, foram

excluídos os pesquisadores que não possuem produção bibliográfica em colaboração com outros pesquisadores da rede social estudada.

Para a obtenção da evolução da rede de cooperação nos períodos de interesse, efetuamos o cálculo das métricas de rede relacionadas com a estrutura: densidade, diâmetro, grau médio e grau médio ponderado (Freeman, 2004), número de componentes, coeficiente de aglomeração e coeficiente de aglomeração médio, comprimento médio de caminho e o comportamento do componente gigante (Newman, 2001c), em termos de quantidade de nós e percentual em relação à rede (Scott, 2002), conforme apresentado na Tabela 01.

Tabela 01 – Escopo da rede – IFS 2008-2018 – Biênios de Interesse

Métrica	2008-2010	2010-2012	2012-2014	2014-2016	2016-2018	2008-2018
Nós	91	119	165	162	162	275
Arestas	69	94	153	153	162	343
Grau médio	1,516	1,58	1,855	1,889	2	2,495
Grau médio ponderado	4,374	4,992	4,8	5,605	6,444	8,771
Comprimento médio de caminho	1,434	1,712	3,714	3,349	2,746	8,465
Componentes (Nós)	33	37	39	40	38	27
Componente Gigante – percentual	(8 nós) 8,79 %	(8 nós) 6,72 %	(29 nós) 17,58%	(29 nós) 17,9%	(21 nós) 12,96%	(214 nós) 77,82%
Diâmetro	3	5	10	9	7	22
Densidade	0,017	0,013	0,011	0,012	0,012	0,009
Coeficiente de aglomeração	0,216	0,223	0,246	0,207	0,236	0,231
Coeficiente de aglomeração médio	0,657	0,522	0,478	0,454	0,479	0,389

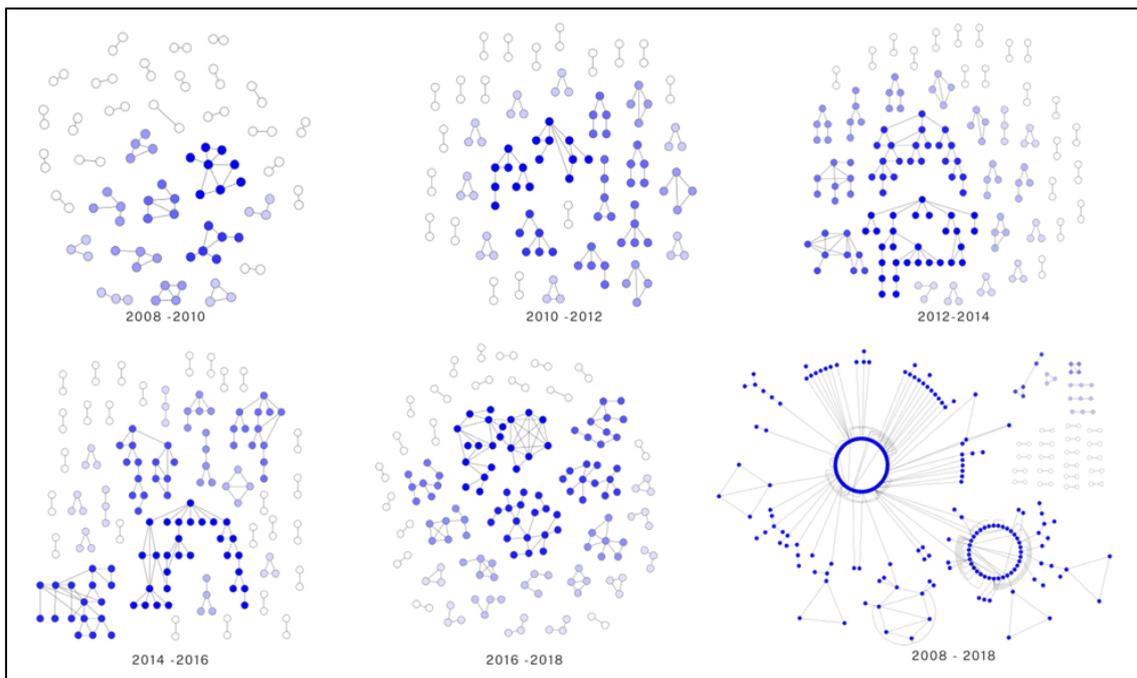
Fonte: Elaborado pelos autores.

A densidade da rede é dada em porcentagem, indicando a relação entre a quantidade total de ligações presentes e a quantidade de ligações possíveis, ou seja, é uma medida do grau de conectividade que determinada rede possui. Descreve o nível que está sendo atingido no momento de ligações potenciais existentes, informando qual o índice de conectividade potencial da rede (Scott; Carrington, 2002).

O diâmetro é uma métrica de rede que diz respeito à maior distância entre os nós de uma rede. No caso em análise, como a rede não é totalmente conectada, quando há vértices desconectados, a distância aferida para atingir nós desconectados é infinita; assim, nossa análise considera a medida do maior componente conectado ou componente gigante. Em redes de coautoria o diâmetro indica a possibilidade de colaboração mais distante, quanto maior o diâmetro, maior o grau de separação entre dois pesquisadores.

A partir dos resultados apresentados na Tabela 01, que apresenta o comportamento estrutural da rede de colaboração estudada, podemos verificar que há uma tendência de crescimento da rede colaborativa entre 2008 e 2014, vide ainda Figuras 08 e 09, com estabilidade a partir do ano de 2014 até o ano de 2018.

Figura 08 – Grafo da evolução temporal da rede de cooperação



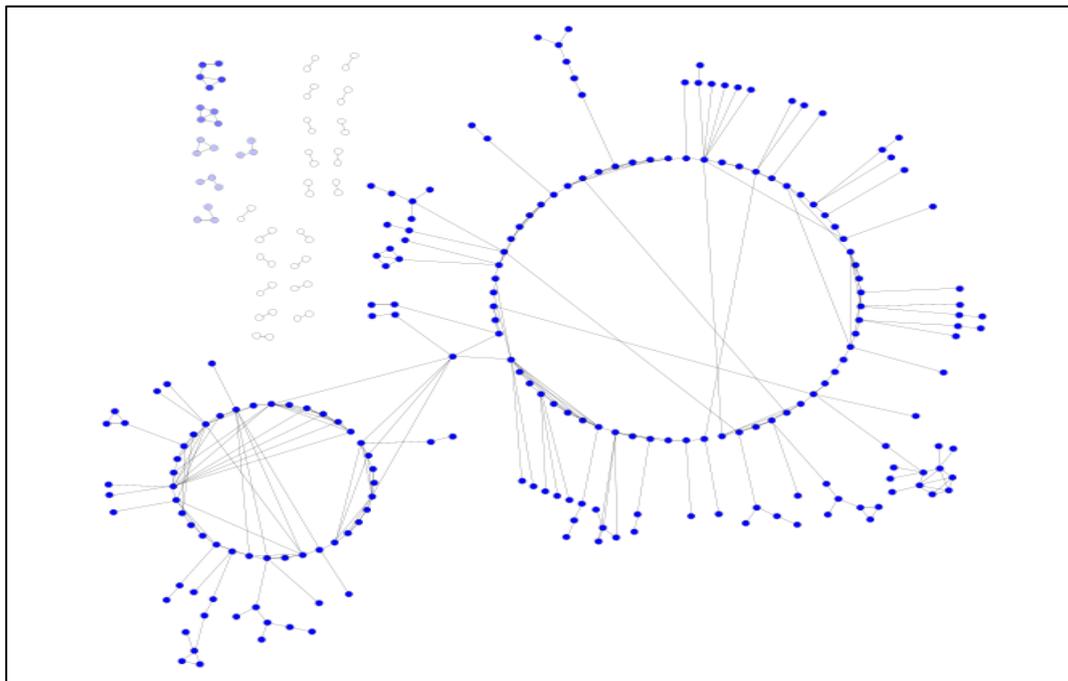
Fonte: Elaborado pelos autores.

Ao maior componente conectado de uma rede denomina-se componente gigante e suas características indicam qual o potencial de colaboração. Quanto à nossa análise, este item de interesse apresenta uma tendência de aumento do percentual de representatividade com relação à rede; entretanto, mantendo-se em um valor pequeno, tendo seu máximo de representatividade nos períodos de

2012 a 2016, vide Tabela 01, contendo aproximadamente 17% dos nós da rede nos dois períodos.

Após as medições e a discussão acerca de escopo e características mais gerais da rede em estudo buscou-se identificar quais os atores mais centrais à rede, utilizando as três medidas de centralidade indicadas por Freeman (1979): centralidade de grau (*degree centrality*), centralidade de intermediação (*betweenness centrality*) e centralidade de proximidade (*closeness centrality*), indicadores de importância relativa e de poder, bem como de controle de fluxo informacional que os pesquisadores podem exercer a partir de sua posição relativa na rede social (Newman, 2001c; Plobación *et al.*, 2009).

Figura 09 – Rede social de colaboração em pesquisa - 2008-2018



Fonte: Elaborado pelos autores.

Devido às características matemáticas e aos limites associados a medição de centralidade de rede, que requerem conexão da rede, aplicamos a estratégia de verificação no componente gigante da rede, onde estão presentes 214 pesquisadores conectados por 304 arestas, ou o equivalente a 77,82 % dos nós e 88,6% das arestas da rede de cooperação completa, conforme dados disponíveis na Tabela 02.

Tabela 02 – Características gerais da Rede de Colaboração

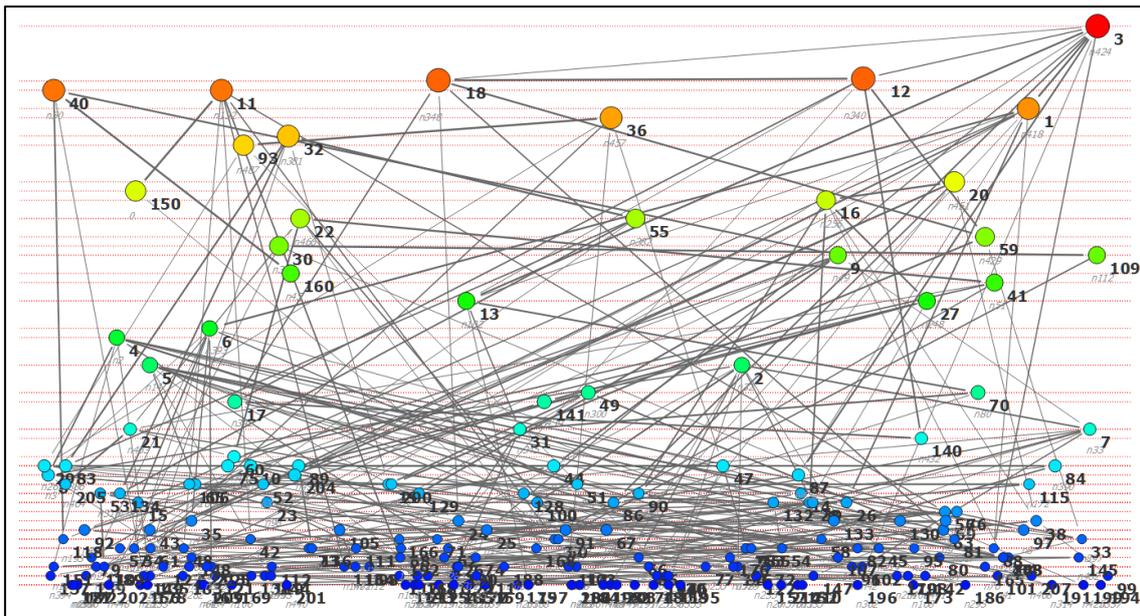
Medida	Valor
Vértices	214
Vértices / Total da Rede	214 de 272 – 77,82%
Arestas	304
Maior distância	22
Distância Média	8,48
Maior grau	12
Menor grau	1
Grau médio	2,841
Maior grau ponderado	62
Menor grau ponderado	1
Grau médio ponderado	10,009
Centralidade de grau	1 - 12
Centralidade de proximidade	0.074 - 0.173
Centralidade de intermediação	0.000 – 10777.328

Fonte: Elaborado pelos autores.

O cálculo e a análise da medida de centralidade de grau permitiram identificar quais pesquisadores possuem a maior quantidade de ligações com outros, ou seja, quais foram os pesquisadores mais colaborativos. A Figura 10 ilustra estas relações através da centralidade de grau da rede, com a distribuição dos elementos da rede em níveis, do maior para o menor, de cima para baixo, de acordo com seu grau e a escala de distribuição. Nesta visualização, consideramos os pesos das arestas para a extração das medidas, definidos pelo grau de colaboração.

A distribuição de grau da rede obedece a uma lei de potência, o que indica uma confirmação dos preceitos indicados por Newman (2001c), com relação a redes de coautoria, afirmando que muitos autores possuem poucas ligações e poucos pesquisadores possuem muitas ligações, indicando ainda a possibilidade da rede ser considerada livre de escala, em que Hubs (Sousa *et al.*, 2019) e agentes centralizadores estão presentes, vide Figura 10.

Figura 10 – Centralidade de Grau – Distribuição por níveis



Fonte: Elaborado pelos autores.

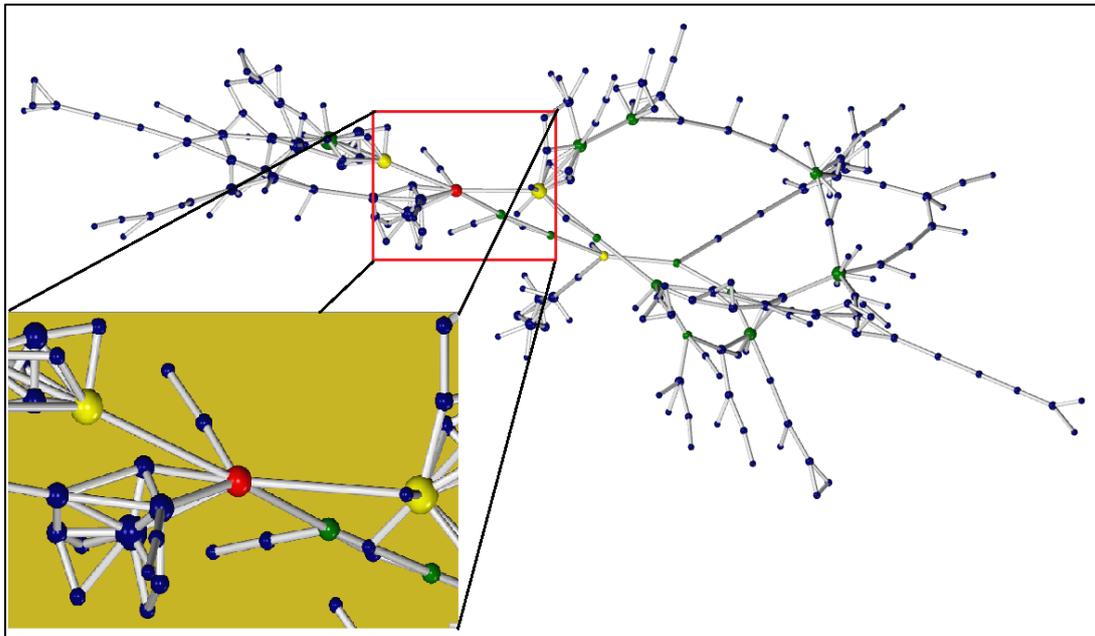
A centralidade de intermediação ou *betweenness*, indica em termos de estrutura da rede quais atores mantêm a coesão da rede ao mesmo tempo em que podem controlar o fluxo de informações entre os conjuntos de pesquisadores não adjacentes. É calculada a partir do caminho geodésico entre o nó alvo e todos os outros nós da rede, identificando quantas geodésicas passam pelo nó em questão. Esta posição ocupada pelo pesquisador o concede o poder de atuar como intermediário no fluxo de informações ou de recursos que trafeguem entre os agentes da rede.

O cálculo é efetuado a partir da construção da matriz de distância ou do comprimento do caminho mais curto (*Shortest Path Length*) da rede analisada, que é uma matriz do tipo $N \times N$, na qual cada elemento (i, j) representa a distância geodésica entre o elemento i e o elemento j .

A partir dos dados da matriz de distância construímos a visualização da rede de pesquisadores apresentada na Figura 11. Destacamos nessa

visualização a presença do pesquisador com maior valor de intermediação (*betweenness*), que possui uma importância relativamente grande para a rede, considerando que sua ausência causaria a desconexão do componente gigante, cuja estrutura é claramente formada por dois grandes *clusters*.

Figura 11 – Componente Gigante - Destaque do pesquisador com maior grau de centralidade de intermediação em vermelho



Fonte: Elaborado pelos autores.

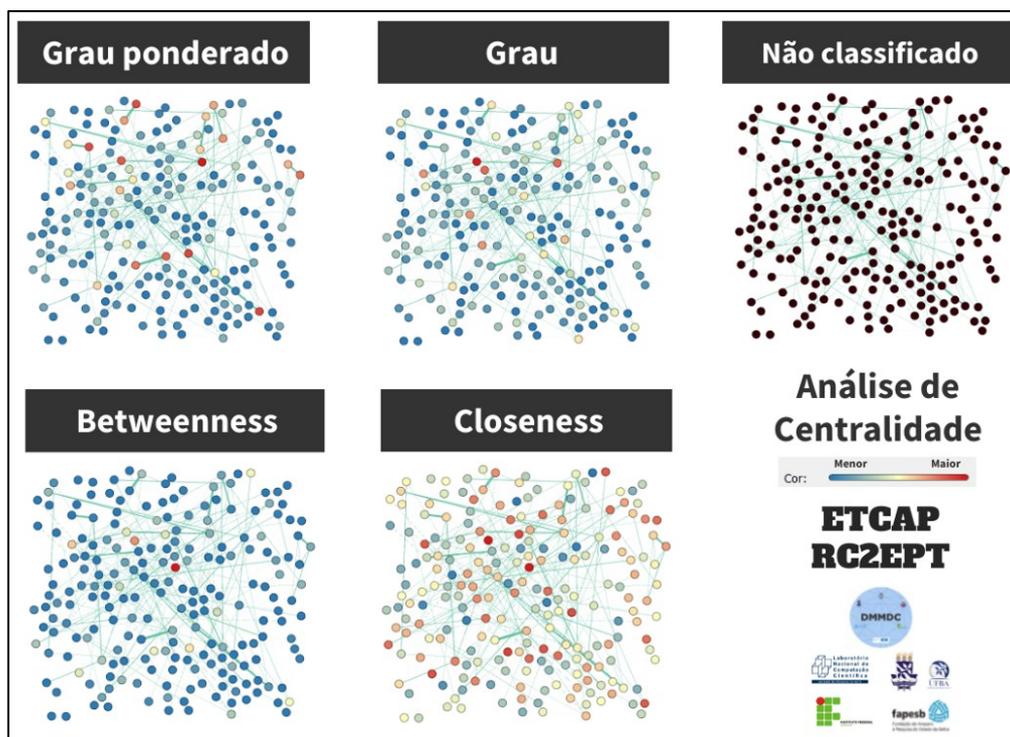
A posição relativa do pesquisador que possui o maior grau de intermediação o classifica como o principal broker (Sousa *et al.*, 2019), cuja ausência ou eliminação causaria a divisão e o aumento de componentes da rede e sua posição relativa na rede o permite atuar como uma ponte entre os dois subgrupos, com potencial para efetuar o controle sobre a circulação de informações entre os elementos, traduzindo sua importância para a rede, tanto com sua ausência quanto com sua presença.

A obtenção da Centralidade de Proximidade se deu a partir do cálculo do valor inverso da soma das distâncias geodésicas dos outros nós da rede, de forma que os mais centrais possuem caminhos mais curtos para alcançar seus pares, podendo exercer influência sobre eles e assim, para toda a rede. A medição da centralidade de proximidade indica novamente o destaque para o pesquisador que além de possuir o maior valor de intermediação (*betweenness*),

também apresenta o maior valor da centralidade de proximidade (*closeness*) da rede estudada, consolidando sua importância e destaque na estrutura.

A visualização é um elemento importante para a Análise de Redes Sociais, pois permite destacar os agentes de interesse, mediante o uso de algoritmos e estratégias que apoiam o pesquisador em sua tarefa de analisar os achados de sua pesquisa, a Figura 12 apresenta de maneira comparativa as medidas de centralidade da rede.

Figura 12 – Visualização comparativa das medidas de centralidade



Fonte: Elaborado pelos autores.

Assim, consideramos importante para a difusão do conhecimento o acesso mais amplo possível aos resultados obtidos após a execução do framework, bem como os programas específicos estão disponíveis ao público, para um acesso interativo e imersivo nas redes e dados estatísticos, mediante acesso identificado, no portal RC2EPT - Portal de Redes de Cooperação Científica na Educação Profissional e Tecnológica.

4 CONCLUSÕES

Neste artigo apresentamos um framework modelado para a construção de modelos de redes de cooperação, com a possibilidade de uso e validação/aplicação em dados empíricos obtidos a partir da base de currículos Lattes, para o domínio da educação profissional e tecnológica.

Esta proposta de framework para um processo de descoberta de conhecimento a partir de bases de informações bibliográficas permite o tratamento e análise de redes de cooperação científicas, identificando como pesquisadores e cientistas atuaram colaborativamente.

A adoção da modelagem científica é base da estratégia de concepção e sua representação é conduzida a partir de um fluxo de processos, associado ao uso de sistemas computacionais para processamento dos dados.

A validação empírica do framework colimou com a disponibilização de uma sistemática de indicadores de produção e de cooperação científica, associada com a obtenção das medidas de centralidade da rede de cooperação do Instituto Federal de Sergipe - IFS, identificando os atores mais centrais da rede, desnudando aqueles que podem exercer influência toda a rede, o que permitiu a visualização dos indicadores de produção e de cooperação científica, a criação de um banco de dados sobre a produtividade acadêmica, todos apresentados de forma interativa e publicados em um portal temático na internet

O presente framework produz uma representação simbólica do processo de descoberta de conhecimento, permitindo difundir os resultados da pesquisa, considerando as restrições e o comportamento do fenômeno observado, possibilitando uma maior compreensão do mesmo com o uso dos recursos disponibilizados pela computação, em um sistema de publicação dos dados na internet denominado de Portal RC2EPT – Redes de Cooperação Científica na Educação Profissional e Tecnológica.

Considerando a possibilidade da apropriação, adaptação e reuso do framework para usos em domínios diversos, interessados poderão se apropriar deste framework e do conjunto de soluções técnicas e programas, atuando na modelagem e execução de processos de descoberta de conhecimento e apoio

a decisão, ampliando suas possibilidades de ação, contribuindo para sua consolidação, pelo uso em diferentes domínios.

Após a validação empírica pelo estudo de caso, consideramos que o framework aqui desenvolvido para modelagem de redes de colaboração científica pode apoiar a ação de pesquisadores da temática de redes de cooperação científica, com uma abordagem multidisciplinar para a concepção dos modelos, que permitirá a adaptação ao contexto das pesquisas em domínios diversos.

Com este modelo ampliam-se as possibilidades de ação de pesquisadores de diversas áreas e campos, que passam a ter disponível um arcabouço teórico-metodológico-prático, adaptável e com uma curva de aprendizado suave, diminuindo a dificuldade de condução de estudos com modelagem de redes de cooperação científica e, ainda, ampliando o arcabouço técnico e processual para a descoberta, gestão e difusão do conhecimento científico.

REFERÊNCIAS

ANDRADE, M. T. T.; ROCHA, E. C.; RIBEIRO, N. M.; PEREIRA, H. B. B. A colaboração em comunidades científicas interdisciplinares: análise de projetos de pesquisa com auxílio da teoria de redes. **Revista Educação, Tecnologia e Cultura - E.T.C.**, [S. l.], n. 11, jul., 2016. Disponível em: <https://11nq.com/4sMfl>. Acesso em: 25 set. 2024.

ARIA, M.; CUCCURULLO, C. Bibliometrix: an r-tool for comprehensive science mapping analysis. **Journal of Informetrics**, [S. l.], v. 11, n. 4, p. 959–975, nov. 2017. Disponível em: <https://www.sciencedirect.com/science/article/pii/S1751157717300500>. Acesso em: 25 set. 2024.

BOAVENTURA NETTO, P. O. JURKIEWICZ, S. **Grafos: introdução e prática**. 2. ed. São Paulo: Blucher, 2017.

BURNHAM, T. F. DMMDC: uma proposta plural de difusão do conhecimento. *In*: MATTA, A. E. R. ROCHA, J. C. (org.). **Cognição: Aspectos contemporâneos da construção e difusão do conhecimento**. Salvador: EDUNEB, 2016. p. 19-46.

DE SOLLA PRICE, D. J. Networks of Scientific Papers. **Science**, [S. l.], v. 149, n. 3683, p. 510–515, 1965. Disponível em:

<https://garfield.library.upenn.edu/papers/pricenetworks1965.pdf>. Acesso em: 25 set. 2024.

DIAS, A.; RUTHES, S.; LIMA, L.; CAMPRA, E.; SILVA, M.; SOUSA, M. B.; PORTO, G. Network centrality analysis in management and accounting sciences. **RAUSP Manag. J.**, [S. l.], v. 55, n. 2, p. 207-226, July, 2020. Disponível em: <https://doi.org/10.1108/rausp-02-2019-0021>. Acesso em: 02 out. 2024.

DIGIAMPIETRI, I. A.; MENA-CHALCO, J. P.; PÉREZ-ALCÁZAR, J. J.; TUESTA, E. F.; DELGADO, K. V.; MUGNAINI, R.; SILVA, G. S. Minerando e Caracterizando Dados de Currículos Lattes. *In*: BRAZILIAN WORKSHOP ON SOCIAL NETWORK ANALYSIS AND MINING (BRASNAM), 2012, Curitiba. **Anais [...]**. Porto Alegre: Sociedade Brasileira de Computação, 2012. p. 117-128. Disponível em: <https://sol.sbc.org.br/index.php/brasnam/article/view/6868>. Acesso em: 02 out. 2024.

FIALHO, J. M. R. A Sociedade das Redes. *In* FIALHO, J. M. R (org.). **Redes Sociais**. Como compreendê-las? – Uma introdução à análise de redes sociais. Lisboa: Edições Silabo, 2020.

FREEMAN, L. C. Centrality in social networks conceptual clarification. **Social Networks**, [S. l.], v. 1, n. 3, 1978–1979, p. 215-239. Disponível em: https://www.researchgate.net/publication/222447730_Centrality_in_Social_Networks'_Conceptual_Clarification. Acesso em: 08 Jan. 2019.

FREEMAN, L. **The development of social network analysis**: a study in the sociology of science. Vancouver: Empirical Press, 2004.

FREITAS, L. Q. **Medidas de centralidade em grafos**. 2010. 111 f. Dissertação (Mestrado em Engenharia de Produção) – Instituto Alberto Luiz Coimbra de Pós-Graduação e Pesquisa (COPPE), Universidade Federal do Rio de Janeiro (UFRJ), Rio de Janeiro, 2010.

GALEFFI, D. A. Ética, privacidade e confidencialidade de informação em saúde: investigando a ética na sociedade do conhecimento, da informação, da aprendizagem e do controle a partir de uma Teorização Polilógica. **Informação em Pauta**, Fortaleza, v. 5, n. especial, p. 9-22, 2020. Disponível em: https://repositorio.ufc.br/bitstream/riufc/50957/1/2020_art_dagaleffi.pdf. Acesso em: 25 set. 2024.

GOMES, C. H. P., CAVALCANTE, P. S., & LINS, W. C. B. Análise de redes de colaboração científica dos docentes do programa de pós-graduação em educação matemática e tecnológica da Universidade Federal de Pernambuco. **Em Teia**: Revista de Educação Matemática e Tecnológica Iberoamericana, Recife, v. 7, n. 2. 2016. Disponível em:

<https://periodicos.ufpe.br/revistas/index.php/emteia/article/view/2784>. Acesso em: 25 set. 2024.

HANNEMAN, R. A.; RIDDLE, M. **Introduction to social network methods**. Riverside, CA: University of California, Riverside, 2005.

HIGGINS, S. S.; RIBEIRO, A. C. R. **Análise de redes em Ciências Sociais**. Brasília: ENAP, 2018.

KARAM-FILHO, J. Princípios Básicos de Modelagem. *In*: MESSEDER, S. A.; CAMBUI, E. C. B. (org.). **Analista cognitivo: uma profissão interdisciplinar**. Salvador: EDUFBA, 2019.

KARAM-FILHO, J.; ALMEIDA, R. C. **Introdução à Modelagem Matemática**. Notas de Aula – Curso de Nivelamento do Programa de Doutorado em Modelagem Computacional do Laboratório Nacional de Computação Científica. LNCC, 2003. 70 p.

LEITE, D.; CAREGNATO, C. E.; LIMA, E. G. S.; PINHO, I.; MIORANDO, B. S.; SILVEIRA, P. B. Avaliação de redes de pesquisa e colaboração. **Avaliação: Revista da Avaliação da Educação Superior, Sorocaba**, v. 19, n. 1, p. 291-312, mar. 2014. Disponível em: <https://www.scielo.br/j/aval/a/tvqZPTRfdvFZBmGrsn7HKhz/>. Acesso em: 25 set. 2024.

LEMIEUX, V.; OUIMET, M. **Análise estrutural das Redes Sociais**. 2. ed. Lisboa. Instituto Piaget. 2012.

MENA-CHALCO, J. P.; CESAR JUNIOR, R. M. ScriptLattes: an open-source knowledge extraction system from the Lattes platform. **Journal of the Brazilian Computer Society**, Porto Alegre, v. 15, p. 31-39, 2009. Disponível em: <https://www.scielo.br/j/jbcos/a/DNqND3wQHrTHkCNWQbKx6pt/?lang=en>. Acesso em: 25 set. 2024.

NEWMAN, M. E. J. Scientific collaboration networks. I. Network construction and fundamental results. **Physical Review E**, [S. l.], v. 64, n. 1, p. 016131, 28 jun. 2001a. Disponível em: <https://periodicos.ufpe.br/revistas/index.php/emteia/article/view/2784>. Acesso em: 25 set. 2024a.

NEWMAN, M. E. J. Scientific collaboration networks. II. Shortest paths, weighted networks, and centrality. **Physical Review E**, [S. l.], v. 64, n. 1, p. 016132, 28 jun. 2001b. Disponível em: <https://periodicos.ufpe.br/revistas/index.php/emteia/article/view/2784>. Acesso em: 25 set. 2024b.

NEWMAN, M. E. J. The structure of scientific collaboration networks. **Proceedings of the National Academy of Sciences**, [S. l.], v. 98, n. 2, p. 404-409, 2001c. Disponível em:

<https://www.pnas.org/doi/abs/10.1073/pnas.98.2.404>. Acesso em: 02 out. 2024c.

POBLACIÓN, D. A.; MUGNAINI, R.; RAMOS, L. M. C. **Redes sociais e colaborativas em informação científica**. São Paulo: Editora Angellara, 2009.

SCOTT, J.; CARRINGTON, P. J. (org.). **The SAGE handbook of social network analysis**. London: SAGE, 2011.

SNIJDERS, T. A. B. Statistical Models for Social Networks. **Annual Review of Sociology**, [S. l.], v. 37, n. 1, p. 131–153, 2011. Disponível em: <https://periodicos.ufpe.br/revistas/index.php/emteia/article/view/2784>. Acesso em: 25 set. 2024.

SOUZA, A. C.; SAMPAIO, R. L.; MACEDO, M. C. F.; FONSECA, P. V. O.; JESUS, V. S.; LIMA, I. A. **Análise de redes sociais: uma abordagem prática**. Salvador: EDUFBA, 2019.

SCIENTIFIC MODELING AND APPLICATION OF A FRAMEWORK FOR ANALYSIS OF SCIENTIFIC COOPERATION NETWORKS

ABSTRACT

Objective: This article aims to present a conceptual/operational framework proposal for knowledge discovery through scientific modeling of scientific cooperation networks.

Methodology: A case study was conducted in the domain of professional and technological education to validate the proposed framework. The process involved data extraction, processing, loading, and analysis, as well as the publication of results with broad access to information, characteristics, indicators, and metrics of the target collaboration network. **Results:** The framework was validated through a case study in the domain of professional and technological education, proving effective in data extraction, processing, loading, and analysis, enabling the publication of results in the domain of interest with broad access to information, characteristics, indicators, and metrics of the target collaboration network. **Conclusions:** Scientific modeling of scientific cooperation networks, using the proposed framework, is a promising strategy for knowledge discovery in this context. The application of social network analysis techniques and the use of information technology and communication solutions allow for efficient exploration of relationships and statistical indicators of collaboration networks.

Descriptors: Scientific Modeling. Scientific Collaboration. Social Networks. Professional and Technological Education.

MODELADO CIENTÍFICO Y APLICACIÓN DE UN FRAMEWORK PARA EL ANÁLISIS DE REDES DE

COOPERACIÓN CIENTÍFICA

RESUMEN

Objetivo: Este artículo tiene como objetivo presentar una propuesta de framework conceptual/operativo para el descubrimiento de conocimiento a través del modelado científico de redes de cooperación científica. **Metodología:** Se utilizó un estudio de caso en el ámbito de la educación profesional y tecnológica para validar el framework propuesto. El proceso involucró la extracción, el tratamiento, la carga y el análisis de los datos, así como la publicación de los resultados con amplio acceso a la información, características, indicadores y métricas de la red de colaboración objetivo. **Resultados:** El framework fue validado a través de un estudio de caso realizado en el ámbito de la educación profesional y tecnológica, demostrando ser efectivo para la extracción, el tratamiento, la carga y el análisis de los datos, permitiendo la publicación de resultados del dominio de interés, con amplio acceso a la información, características, indicadores y métricas de la red de colaboración objetivo. **Conclusiones:** El modelado científico de redes de cooperación científica, utilizando el *framework* propuesto, es una estrategia prometedora para el descubrimiento de conocimiento en este contexto. La aplicación de técnicas de análisis de redes sociales y el uso de soluciones de tecnología de la información y comunicación permiten explorar de manera eficiente las relaciones y los indicadores estadísticos de las redes de colaboración.

Descriptores: Modelado Científico. Colaboración Científica. Redes Sociales. Educación Profesional y Tecnológica.

Recibido em: 30.11.2023

Aceito em: 10.09.2024