

O ACERVO DA TEIAS EM FOCO: as perspectivas e contribuições da revista através de meta-análises

*Gustavo Silva Semaan
Edelberto Franco Silva
José André de Moura Brito*

Resumo

Com o objetivo de conhecer novas perspectivas em relação ao acervo da Revista TEIAS, em destaque pela alta qualidade de seus trabalhos e de excelência confirmada pelo sistema de avaliação Qualis da CAPES, o trabalho analisa sua comunidade por meio das redes de colaborações e contribuições. Criada no ano 2000 com foco na área de conhecimento Educação, em seu 21º ano de atuação superou a marca de 1.000 artigos publicados por mais de 1.300 autores. Para a realização das meta-análises foi considerado o processo de Descoberta de Conhecimento em Bases de Dados e conceitos de Teoria dos Grafos. Os resultados obtidos possuem estatísticas gerais, análises das redes de colaboração, termos em destaque ao passar dos anos e autores considerados influentes com base em sua frequência de publicação na revista, pela quantidade de autores com quem colaboram bem como por meio do uso da medida de centralidade por intermediação. Assim, o presente trabalho fornece informações e análises importantes, que podem ser consideradas em outras pesquisas e apoiam interpretações, perspectivas e contribuições deste importante acervo.

Palavras-chave: educação; grafos; meta-análise; KDD; medidas de centralidade, redes de colaboração.

THE TEIAS COLLECTION IN FOCUS: perspectives and contributions through meta-analysis

Abstract

Aiming to discover new perspectives about the 'TEIAS' collection, highlighted by the high quality and excellence confirmed by the CAPES Qualis evaluation system, the work analyzes its community through the networks of collaborations and contributions. Created in 2000 with a focus on Education, in its 21st year of operation it transcended the mark of 1,000 articles published by more than 1,300 authors. To carry out the meta-analyses, the Knowledge Discovery in Databases process and concepts of Graph Theory were considered. The results obtained bring general statistics, analyses of the collaboration networks, highlighted terms over the years, and authors considered influential, based on their frequency of publications, the number of authors with whom they collaborate as well as through the use of between centrality measure. In this sense, the present work provides solid information, which can be regarded in other investigations and supports interpretations, perspectives, and contributions of this important collection.

Keywords: education; graph theory; meta-analysis; KDD; centrality measures; collaboration network.

EL ACERVO DE TEIAS EN ENFOQUE: perspectivas y contribuciones a través del metaanálisis

Resumen

Con el objetivo de conocer nuevas perspectivas en relación con la colección de la revista TEIAS, destacada por la alta calidad y excelencia de sus trabajos confirmada por el sistema de evaluación Qualis de la CAPES, el trabajo analiza su comunidad a través de las redes de colaboraciones y contribuciones. Creada en el año 2,000 con un enfoque en el área de conocimiento de la Educación, en su 21º año de operación superó la marca de 1,000 artículos publicados por más de 1,300 autores. Para la realización de los metaanálisis fue considerado el proceso de Descubrimiento del Conocimiento en Bases de Datos y conceptos de la Teoría de Grafos. Los resultados obtenidos tienen estadísticas generales, análisis de las redes de colaboración, términos destacados a lo largo de los años y autores considerados influyentes en función de su frecuencia de publicación en la revista, por la cantidad de autores con los que colaboran, bien como por medio del uso de la medida de centralidad a través de la intermediación. De esta manera, el presente trabajo proporciona informaciones sólidas, que pueden ser consideradas en otras investigaciones y apoyan interpretaciones, perspectivas y contribuciones de esta importante colección.

Palabras clave: educación; teoría de grafos; metaanálisis; KDD; medidas de centralidad.

INTRODUÇÃO

O presente trabalho tem como objetivo principal fornecer perspectivas em relação às contribuições e colaborações do acervo da revista TEIAS. O conhecimento é construído a partir troca de experiências e colaborações que, nem sempre, são formalizadas em produtos como artigos científicos ou livros. De fato, efetivamente, por meio de técnicas de meta-análises é possível identificá-las principalmente através das relações de coautorias nos artigos ou mesmo em suas referências. Estudos sobre comunidades em áreas específicas, colaborações e acervos são alvo de intensa pesquisa em diversas áreas (OLIVEIRA, 2012; SEMAAN *et al.*, 2019a, 2019b).

A Revista Teias tem como escopo “[...] o incentivo à pesquisa e ao debate da área da educação, centrados nas relações com a sociedade [...]”, e contribui fortemente com a democratização do conhecimento por meio da disponibilização livre, imediata e gratuita de seu acervo, utilizando-se de diversos indexadores nacionais e internacionais, além das bibliotecas virtuais que disponibilizam o sistema OAI (do inglês Open Archives Initiative) para consultas (TEIAS, 2020).

Criada no ano de 2000, com foco na área de conhecimento Educação, em 2020 a revista superou a marca de 1.000 artigos publicados por mais de 1.300 autores em seu acervo, que possui grande relevância e destaque, com excelência comprovada, inclusive, pelo sistema de avaliação Qualis da CAPES (Coordenação de Aperfeiçoamento de Pessoal de Nível Superior) (TEIAS, 2020).

Perspectivas por meta-análises

Para a obtenção de perspectivas através dos metadados de acervos, comumente são consideradas as seguintes dimensões: (i) relacionadas ao tempo (ano, edição e/ou número da publicação); (ii) análise do título e/ou das palavras-chave; (iii) as relações entre coautores de artigos; e (iv) as citações relacionadas dos artigos. A combinação de tais atributos pode produzir

diferentes panoramas de um mesmo acervo, de acordo com a maneira em que os dados são trabalhados ou conforme o ponto de vista de quem analisa os resultados.

O objetivo do presente trabalho é apresentar perspectivas e contribuições da Revista TEIAS com base apenas nos metadados de seu acervo, com análises objetivas e desprovidas de intencionalidade política, restritas aos dados coletados em seu portal, que não abordam a processualidade e qualidade das publicações. Para isso são utilizados conceitos de Teoria dos Grafos (BOAVENTURA, 2001), o processo de KDD (Descoberta de conhecimento em bases de dados, do inglês *Knowledge Discovery in Databases*) (GOLDSCHMIDT, PASSOS, 2005) e obtenção de estatísticas gerais.

Ainda em relação ao contexto de meta-análises, em geral, as pesquisas têm como foco a busca por panoramas em relação a acervos específicos, sejam de importantes revistas ou mesmo de eventos científicos (como congressos e simpósios). Entretanto, em outra direção, métodos como revisões sistemáticas da literatura podem ser considerados para pesquisas sobre temas específicos, e assim possibilitam contemplar diversos repositórios de dados simultaneamente.

Objetivos específicos

Para alcançar o objetivo geral e fornecer perspectivas em relação às características, contribuições e colaborações da Revista TEIAS, que possui grande destaque e importância no cenário nacional, foram organizados como objetivos específicos: (i) Coleta e pré-processamento conforme o KDD: recuperar, processar, organizar, enriquecer, unificar e consolidar dados do acervo para compor uma base de informações consistente; (ii) Estatísticas gerais: mediante a análise dos trabalhos e dos metadados disponíveis; (iii) Termos: os títulos dos trabalhos foram analisados por meio da etapa de mineração de dados, mais especificamente, da mineração de texto (do inglês *Text Mining*), do processo de KDD; (iv) Autoria e colaborações: após identificação e unificação dos autores, é possível obter a quantidade de artigos publicados (frequência) e seus graus de colaboração através das coautorias existentes; (v) Análises de colaboração: por meio de modelagens em grafos e de técnicas de análise em redes de colaboração (do inglês *Collaborations Network*), é possível construir uma “Teia” (um grafo ou uma rede) que ilustra a relação entre os autores; (vi) Construir um contexto contemplando os objetivos específicos, por meio de medidas de centralidade em grafos, a frequência de publicação e as colaborações dos (co)autores; e, por fim, (vii) com base nas análises realizadas, identificar e apresentar uma relação de autores considerados influentes conforme os critérios adotados.

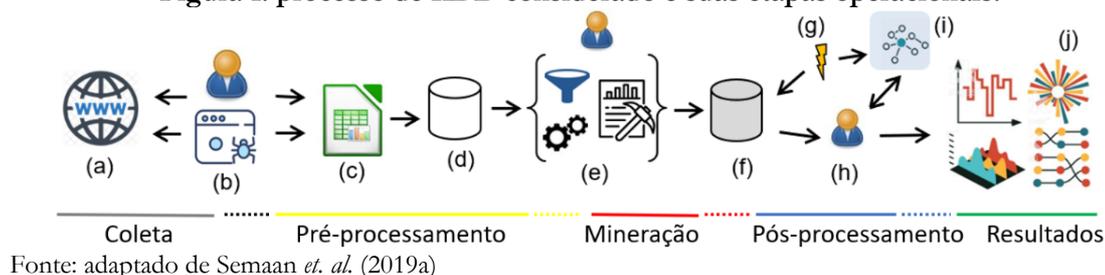
Além da Introdução, o trabalho está organizado em outras 4 seções. A seção *Metodologia* detalha o processo de KDD, desde a obtenção dos dados até a construção dos grafos considerados. A seção *Revisão da literatura e perspectivas* apresenta trabalhos semelhantes relacionados às meta-análises, e estatísticas gerais do acervo da revista bem como termos em destaque nos títulos dos artigos. Na seção *Redes de colaboração* são apresentados os grafos construídos para retratar as redes de colaboração em diferentes *layouts*, conforme as centralidades de grau e de intermediação. Ainda nessa seção, os autores considerados mais influentes são citados o que, de certa maneira, resulta em uma singela homenagem. Por fim, a última seção relata as conclusões, e as propostas para novas pesquisas.

METODOLOGIA

A metodologia utilizada no trabalho, especificamente em relação ao processamento de dados e direcionamento do processo de KKD, é semelhante às propostas recentes da literatura. Para a área *Informática na Educação* SEMAAN *et. al.* (2019a) analisaram o acervo do Simpósio Brasileiro de Informática na Educação (SBIE), enquanto SEMAAN *et. al.* (2019b) e SEMAAN *et. al.* (2019c) apresentaram, respectivamente, um panorama geral com estatísticas e termos de interesse e análises das redes de colaborações do SIMPEP (Simpósio de Engenharia de Produção). Destaca-se que, devido às características específicas do acervo TEIAS e de seu portal, onde os dados foram coletados, ajustes e procedimentos adicionais foram realizados.

De modo objetivo, esta seção apresenta o processo KDD desde a coleta dos metadados e as etapas de pré-processamento para a consolidação dos dados, até a construção dos gráficos e dos grafos de colaborações entre autores. A Figura 1 sintetiza as etapas operacionais do processo de KDD, de modo a ser possível sua reprodução. Destaca-se que, embora o processo seja detalhado, adequações são necessárias conforme os dados de entrada e objetivos específicos a serem alcançados.

Figura 1: processo de KDD considerado e suas etapas operacionais.



Coleta e pré-processamento de dados

Em um primeiro momento, para construir a base de dados, foi necessário obter os metadados dos trabalhos no Portal da Revista (TEIAS, 2020). Nesse sentido, destaca-se o compromisso e a Política de Acesso Livre da revista, com seu acervo disponibilizado de forma imediata, organizada e transparente, contribuindo com a democratização do conhecimento e do acesso à informação.

Com o intuito de ilustrar a ordem de grandeza da comunidade em análise, a base de dados inicial, antes de sua submissão à etapa de pré-processamento do processo de KDD, possuía 56 edições em 21 anos de atuação, com 1.016 artigos publicados por 1.648 autores, unificados em 1.323 após a referida etapa. A meta-análise realizada fornecerá informações sólidas que podem servir de base para trabalhos futuros em diversas vertentes, em especial, para a própria comunidade abordada.

Para a coleta dos metadados dos artigos foi utilizado um aplicativo *webcrawler*, capaz de automatizar parte do processo, e coletas realizadas de maneira manual feitas por pessoas (Figura 1 (b) e (c)). Foram obtidos os dados: ano, edição, título do trabalho e relação de (co)autores. De maneira geral, esses dados foram organizados, modelados e normalizados conforme conceitos de bancos de dados para, posteriormente, iniciar o processo de análise (Figura 1 (d)).

Em posse dos dados, inicia-se a etapa de pré-processamento, em que ocorre a organização, o enriquecimento, a consolidação e a formatação dos dados (Figura 1 (e)). Em geral, um título de um artigo científico busca expressar de maneira objetiva seu propósito. Nesse sentido, foi considerada a tarefa de mineração de texto na etapa de mineração de dados (do inglês *Data Mining*), com o intuito de identificar termos em destaque (GOLDSCHMIDT, PASSOS, 2005).

Mineração de dados

Conforme a Figura 1 (e e f), os títulos são fragmentados em palavras isoladas (atividade denominada *tokenização*) e deve-se eliminar palavras sem relevância semântica (denominadas *stopwords*), como conjunções e preposições. Em seguida é necessário consolidar esses termos, sendo utilizado o processo de unificação por redução ao radical ou equivalência à palavra semelhante mais frequente. Os termos consolidados são inseridos no banco de dados (Figura 1 (f)), mas mantém-se a associação entre os termos consolidados e os autores dos trabalhos em que o termo foi utilizado. Com os dados processados, é possível relacionar suas frequências em relação aos demais termos, em relação aos autores que os utilizam, em função do tempo e de várias outras maneiras, conforme o objetivo da análise (GOLDSCHMIDT, PASSOS, 2005; ARANHA, 2007; PROCACI *et al.*, 2015; SEMAAN *et al.*, 2019a).

Com foco nos nomes dos (co)autores, devido à variedade de maneiras de escrevê-los, incluindo erros de digitação e de grafia ou mesmo abreviações e omissões de parte(s) do sobrenome, foi utilizada verificação humana para decidir se dois nomes correspondem à mesma pessoa (Figura 1 (g) e (h)). Realizada a consolidação de autores, um identificador único é atribuído a cada autor. Destaca-se o caráter iterativo e interativo do KDD, em especial no que concerne à etapa de mineração de dados, em que conforme o resultado obtido pode ser necessário realizar novamente o processamento, com novos ajustes de configurações dos algoritmos. Por isso, ainda na Figura 1 (e), existe um profissional à frente dessa etapa, identificada entre chaves, além do simbólico ícone da mineração: uma picareta.

Modelagens em grafos

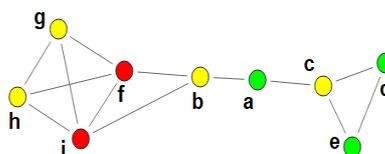
Em trabalhos relacionados às análises de coautorias e colaborações, conceitos relacionados à Teoria dos Grafos comumente são apresentados (PROCACI *et al.*, 2015; MAGALHÃES *et al.*, 2013; NEWMAN, 2001). De maneira formal, porém objetiva, um grafo $G = (V, A)$ é definido como uma estrutura matemática composta de um conjunto de vértices V (pontos) e de arestas A (linhas), em que cada aresta conecta dois vértices. Apesar do rigor científico nas pesquisas e nas análises realizadas, de modo a simplificar a contextualização no trabalho, foram suprimidos alguns formalismos matemáticos e as fórmulas necessários ao estudo, que podem ser obtidos nas referências deste trabalho.

No contexto de redes de colaboração, cada autor é representado por um vértice e, quando dois autores atuaram juntos em um trabalho, uma aresta é adicionada entre eles, indicando a existência de colaboração. Para facilitar o entendimento da modelagem proposta, adotada amplamente na literatura, a Figura 2 apresenta um acervo exemplo com cinco trabalhos e nove coautores, em que a e b são os autores do artigo 1 (Id=1). A mesma figura apresenta um grafo que modela esse conjunto de dados.

A centralidade de grau de um vértice (do inglês, *Degree Centrality*), ou simplesmente seu grau, corresponde à quantidade de arestas incidentes a ele, e pode ser interpretada como a sua probabilidade de receber informações na rede. Assim, no grafo exemplo, o autor *i* possui grau 4 e tem como vértices adjacentes *b*, *f*, *g* e *h* (BEAUCHAMP, 1965; FREEMAN, 1978). No *layout* do grafo ilustrado na Figura 2, as cores dos vértices indicam os seus diferentes graus e, independentemente da quantidade de colaborações entre um mesmo par de coautores, apenas uma aresta é utilizada na representação.

Figura 2: acervo exemplo e seu grafo.

Id	Autores
1	a b
2	a c
3	c d e
4	b f i
5	f g h i



Fonte: Autoria própria

Por fim, a centralidade de intermediação (do inglês *Betweenness Centrality*) supõe que o fluxo de informação na rede ocorra estritamente pelos caminhos mais curtos (nas denominadas distâncias geodésicas). Essa medida de centralidade possibilita identificar vértices importantes em relação ao controle da comunicação. Assim, ainda sobre a Figura 2, a aresta que conecta os autores *a* e *b* é a única forma de conectar qualquer par de vértices entre os vizinhos de *c* (vértices adjacentes *d* e *e*) e os demais vértices do grafo. Portanto, os vértices *a* e *b* são considerados mais influentes nesse cenário (FREEMAN, 1977).

Pós-processamento e resultados

Por fim, através das perspectivas geradas com base nos dados do acervo, análises importantes podem ser realizadas. Por exemplo, com base nas estatísticas gerais, é possível identificar autores que participam de mais edições, que publicam artigos isoladamente (sem coautores) ou que publicam com frequência acima da média. Além disso, outras visualizações podem ser construídas através de Teoria dos Grafos, em que é possível contabilizar as parcerias de um dado autor e sua influência no tráfego de informações pela rede, mas também visualizar graficamente as “Teias” como mapas de colaboração da comunidade. De maneira adicional, os termos em destaque dos títulos podem ser organizados e apresentados em diferentes dimensões, como em função do tempo, de autores ou mesmo a relação entre os próprios termos. A Figura 1 (j) ilustra os resultados do processo de KDD, com diversos relatórios, gráficos, grafos e perspectivas do acervo com o propósito do presente trabalho em relação ao acervo da revista TEIAS.

REVISÃO DA LITERATURA E PERSPECTIVAS

Trabalhos relacionados a meta-análises fornecem panoramas gerais, sejam em relação a acervos de importantes revistas e eventos ou sobre áreas e temas específicos. Para o primeiro caso, parte-se direto para a coleta do(s) acervo(s) alvo(s) e inicia-se a manipulação de dados. Quando a abordagem trata de um tema ou de uma área, metodologias como as revisões sistemáticas da literatura são frequentemente utilizadas. Em comum, para os dois casos, os

metadados correspondem aos dados de entrada. Assim, estatísticas gerais, modelagem em grafos e outras representações são utilizadas com o intuito de fornecer diferentes perspectivas do objeto de avaliação (acervo ou tema). Tais perspectivas, também denominados panoramas na literatura, (do inglês *overview*) podem ser utilizadas como instrumentos, sejam para efetiva tomada de decisão ou mesmo como indicadores importantes, conforme o objetivo almejado.

Trabalhos relevantes na literatura abordam diversas áreas. Em relação a um tema ou sem ter como objetivo um acervo (ou conjunto de acervos) específico(s), Vanz (2009) analisou a colaboração entre autores de artigos brasileiros indexados na base *Web of Science*, com cerca de 50 mil artigos. Com foco na área *Informática na Educação* (IE), Fernandes e Santos (1999) mapearam grupos atuantes com o intuito de estimular o compartilhamento de informações e colaborações entre eles.

Em relação a acervos específicos, Oliveira (2012) e Rodrigues e Ralha (2015) atuaram em metadados para analisar a comunidade de Sistemas de Informação do Simpósio Brasileiro de Sistemas de Informação (SBSI). Cavalcanti e Silva (2011) apresentou análises semelhantes em relação ao Simpósio Brasileiro de Engenharia de Software (SBES), enquanto Lima *et al.* (2017) abordou as redes de colaborações do Simpósio Brasileiro de Banco de Dados (SBBDD). Semaan *et al.* (2019b) e Semaan *et al.* (2019c) consideraram o acervo do Simpósio de Engenharia de Produção (SIMPEP) para fornecer, respectivamente, um panorama geral e análises das redes de colaborações.

Um pouco mais próximo à temática da revista TEIAS, em uma área que envolve Educação, porém, em um contexto mais específico e relacionado ao uso de Tecnologias de Informação (TI) e Tecnologias de Informação e Comunicação (TIC), Magalhães *et al.* (2013), Procaci *et al.* (2015) e Semaan *et al.* (2019a) atuaram com foco no acervo do Simpósio Brasileiro de Informática na Educação (SBIE). Destaca-se que Semaan *et al.* (2019a) abordaram o acervo com maior volume de dados, e contemplaram todo o acervo existente até a publicação do artigo, além de sua metodologia ter sido a base e a inspiração para a presente pesquisa. De maneira adicional, a história da IE no cenário nacional teve início há cerca de 40 anos, através de experiências desenvolvidas em grupos de pesquisas da UFRJ, UFRGS e UNICAMP (ALVES, FLORENTINO, DIAS, 2007). Nesse sentido, em Bottentuit Jr. *et al.* (2012) são apresentadas TICs para promover e desenvolver competências, habilidades e estimular alunos na busca pelo conhecimento.

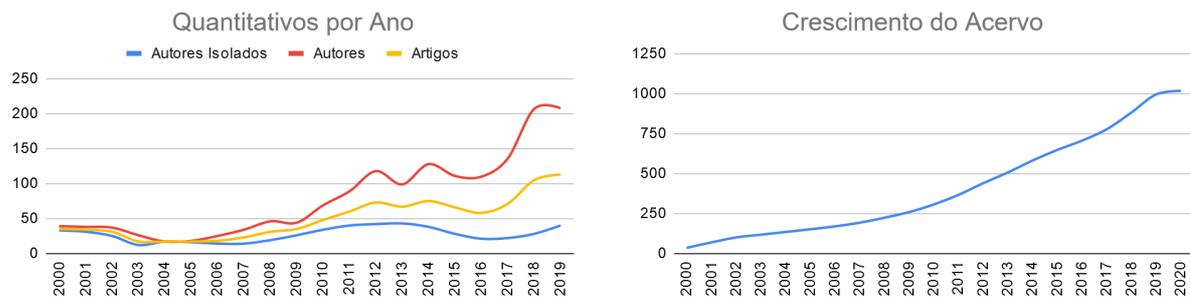
Panorama geral

No momento em que o levantamento foi realizado (TEIAS, 2020), o acervo da revista possuía 1.016 artigos publicados por 1.648 autores, em que 547 trabalhos foram publicados por autores isolados (SA, do inglês *single-author*), o que corresponde a cerca de 54% do acervo. Trata-se de um alto percentual quando comparado aos acervos citados no início da seção, em que as colaborações ocorrem com maior frequência e em maior volume (quantidade de autores em um mesmo artigo). Nesse sentido, deve-se considerar as especificidades das áreas e as características dos processos editoriais. Após a unificação de autores, seu quantitativo foi reduzido para 1.016, e tornou-se possível a realização dos cálculos das estatísticas gerais.

Com o intuito de analisar com rigor os dados dos acervos, e apresentá-los na dimensão tempo, gráficos foram gerados com o período *Ano* especificado no eixo das abscissas. A Figura 3 apresenta os Gráficos *Quantitativos por Ano*, que contemplam os autores isolados, os autores (em geral) e o total de artigos. O outro gráfico da Figura 3 ilustra o *Crescimento do acervo* da revista, que

ultrapassou a marca de mil artigos no início de 2020. Ainda sobre o quantitativo de autores por ano, destaca-se que um mesmo autor pode ser contabilizado mais de uma vez, em anos distintos. Além disso, devido à coleta parcial com apenas de parte de 2020, esse ano foi suprimido no primeiro gráfico para que não fosse transmitida a falsa sensação de queda nos quantitativos.

Figura 3: Gráfico com quantitativos e Gráfico do crescimento do acervo.

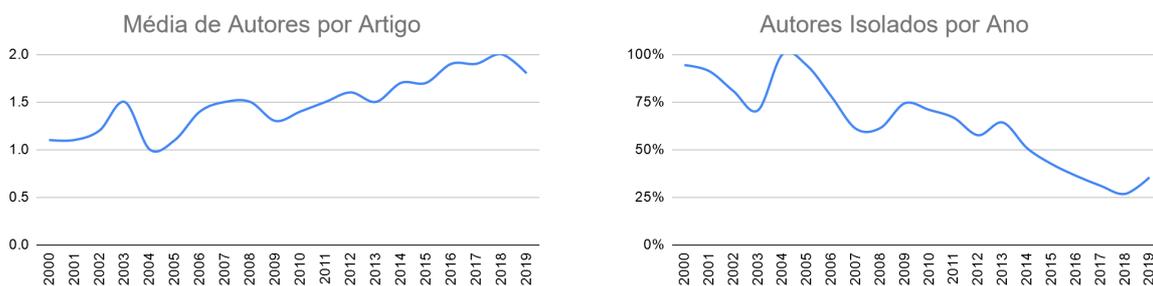


Fonte: Autoria própria

Ao abordar o tema colaboração, com base estritamente nas estatísticas gerais, os dois quantitativos possuem grande relevância, e devem ser analisados: (i) a média de autores por artigo e (ii) o percentual de autores isolados. Com base nos gráficos da Figura 4 observa-se o crescimento da *Média de autores por artigo* ao longo dos anos, de cerca de um no ano 2000 para dois em 2018. Além disso, o gráfico *Autores isolados por ano* indica a redução desse percentual ao longo dos anos, em que em 2000 mais de 94% dos trabalhos eram de SA, em 2004 todos os trabalhos foram de SA, e ao longo do tempo houve queda nos percentuais, chegando a cerca de 25% em 2018. Essas duas análises, em conjunto, são fortes indícios do aumento da colaboração entre pesquisadores.

Destaca-se que ambos os cenários descritos têm como base os metadados de documentos publicados na revista. Entretanto, ressalta-se um componente que não pode ser identificado em meio às metodologias focadas em acervos: as colaborações informais ou não documentadas.

Figura 4: Gráficos com Média de autores por artigo e Autores isolados.



Fonte: Autoria própria

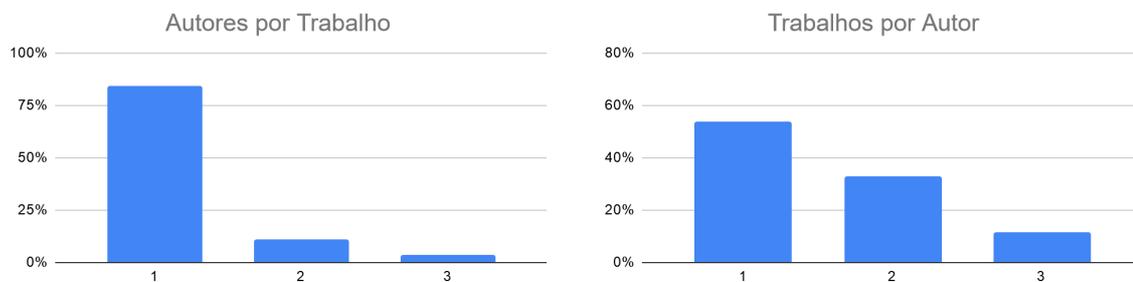
Conforme Oliveira (2012), colaborações entre autores podem ocorrer de diversas maneiras, como em debates acadêmicos, sugestões de revisores em revistas, e comentários e perguntas de participantes de congressos, mas nem sempre resultam em publicações. Nesse sentido, intervenções importantes, como dicas em um encontro informal para um “*breve cafezinho*” podem redirecionar um caminho de pesquisa, mas de fato, frequentemente não são passíveis de

deteção. Trata-se de ponto fraco ao considerar metodologias que atuam com metadados, mas que não diminui as contribuições dos resultados obtidos com tais abordagens.

Ainda em relação à coautoria, conforme o Gráfico *Autores por trabalho* (Figura 5), cerca de 84% dos autores possuem apenas um trabalho no acervo, e mais de 98% deles possuem até 3 artigos. No gráfico *Trabalhos por autor*, cerca de 53,8% trabalharam de maneira isolada, porém 46,2% dos autores atuaram em colaboração, com artigos assinados por dois, três ou quatro coautores.

De modo adicional, existem casos isolados, em que um mesmo autor possui nove publicações no acervo, 14 artigos possuem quatro coautores e quatro artigos possuem seis autores. Tais fatos ocorreram entre os anos 2010 e 2014, exceto em dois artigos que possuem quatro autores, publicados em 2018 na área *Seção Temática* e em 2019 na área *Resenha*. Esses dados foram apresentados, uma vez que o formato atual da revista recomenda o máximo de três coautores por trabalho, e isso tem impacto direto nas colaborações dessa comunidade.

Figura 5: Gráficos Autores por trabalho e Trabalhos por quantidade de autor(es).



Fonte: Autoria própria

Análise de termos de interesse

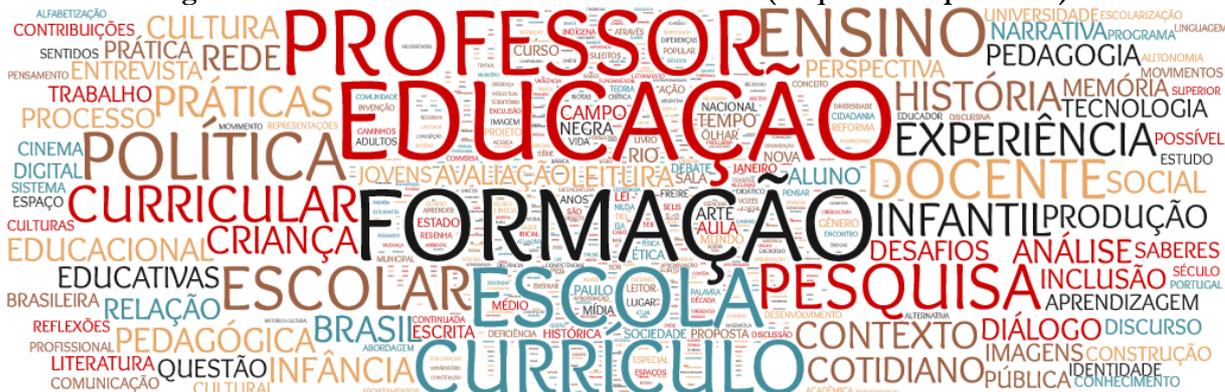
Em relação aos termos em destaque presentes nos títulos dos trabalhos, foram relacionadas as 20 palavras mais frequentes, e seus percentuais de ocorrências foram apresentados por ano na Figura 6. As palavras *Educação*, *Formação* e *Professor* possuíram as maiores médias de ocorrência por ano, com os valores 10,6, 5,5 e 5,1, respectivamente. Além disso, foram analisadas as palavras que ocorreram em mais anos de publicações: *Educação*, *Formação*, *Professor* e *Escolar*, com 21, 20, 19 e 19 ocorrências, respectivamente. Assim, a palavra *Educação*, além de ter sido utilizada em títulos de trabalhos em todos os anos de publicação, possui a maior média de ocorrência.

Figura 6: Percentuais das palavras em destaque nos títulos dos artigos por ano.

	2000	2001	2002	2003	2004	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018	2019	2020
BRASIL	3%	0%	3%	0%	3%	3%	0%	0%	3%	3%	3%	9%	12%	3%	18%	3%	12%	9%	6%	6%	3%
CONTEXTO	0%	3%	3%	0%	9%	6%	9%	0%	3%	0%	6%	6%	3%	3%	3%	9%	3%	3%	9%	26%	0%
CRIANÇA	0%	3%	3%	6%	6%	11%	0%	3%	0%	11%	3%	0%	6%	3%	3%	8%	0%	8%	22%	6%	0%
CURRICULAR	9%	0%	0%	0%	0%	0%	6%	0%	0%	0%	9%	2%	9%	8%	8%	9%	11%	4%	15%	9%	0%
CURRÍCULO	1%	0%	0%	0%	1%	0%	1%	4%	0%	2%	6%	1%	14%	6%	6%	5%	5%	17%	12%	17%	0%
DOCENTE	2%	2%	2%	2%	2%	2%	6%	0%	0%	4%	4%	4%	11%	6%	15%	2%	7%	11%	6%	13%	6%
EDUCAÇÃO	4%	3%	4%	2%	1%	1%	1%	1%	3%	1%	2%	4%	10%	7%	7%	6%	9%	5%	11%	17%	1%
ENSINO	3%	2%	3%	0%	0%	0%	3%	0%	2%	5%	3%	8%	10%	5%	11%	0%	6%	19%	5%	13%	3%
ESCOLA	5%	6%	2%	1%	2%	2%	4%	5%	2%	1%	2%	7%	4%	6%	11%	5%	5%	14%	13%	1%	1%
ESCOLAR	2%	7%	5%	0%	2%	0%	2%	5%	3%	7%	5%	10%	3%	3%	5%	12%	5%	7%	8%	7%	2%
EXPERIÊNCIA	2%	2%	2%	2%	0%	0%	0%	0%	2%	2%	5%	5%	2%	14%	9%	7%	5%	16%	12%	7%	5%
FORMAÇÃO	5%	1%	5%	3%	4%	1%	4%	0%	1%	1%	7%	9%	9%	3%	7%	4%	1%	14%	7%	9%	6%
HISTÓRIA	0%	3%	0%	3%	0%	0%	0%	0%	5%	3%	13%	5%	26%	11%	3%	3%	5%	5%	5%	8%	3%
INFÂNCIA	3%	0%	0%	3%	0%	18%	0%	3%	0%	9%	0%	0%	0%	12%	0%	18%	3%	0%	27%	3%	0%
INFANTIL	0%	0%	2%	0%	0%	14%	2%	2%	0%	2%	0%	2%	2%	2%	5%	14%	2%	5%	23%	16%	5%
PESQUISA	0%	3%	3%	2%	0%	2%	0%	2%	2%	2%	7%	0%	12%	7%	12%	22%	5%	2%	7%	12%	0%
POLÍTICA	2%	3%	5%	0%	2%	0%	2%	0%	3%	2%	8%	6%	5%	5%	8%	5%	8%	11%	19%	6%	0%
PRÁTICAS	3%	3%	3%	0%	0%	3%	8%	5%	0%	3%	3%	15%	10%	3%	20%	3%	5%	10%	5%	0%	3%
PROFESSOR	6%	6%	5%	0%	2%	2%	3%	4%	2%	1%	5%	8%	8%	7%	10%	3%	0%	9%	8%	8%	4%
REDE	3%	0%	0%	0%	0%	0%	0%	0%	0%	3%	6%	0%	12%	6%	6%	9%	12%	6%	6%	21%	9%

Fonte: Autoria própria

Figura 7: Nuvem de Palavras do acervo da Teias (frequência superior a 1).



Fonte: Autoria própria

Uma nuvem de palavras (do inglês *word cloud*) consiste em uma maneira gráfica (ilustrativa) de apresentar a frequência de um conjunto de palavras em um texto ou conjunto de dados. Assim, quanto maior a frequência de utilização de uma palavra, mais chamativa é a sua representação na imagem em relação à disposição e tamanho. Após a unificação dos termos semelhantes, conforme a metodologia citada, além de não considerar as *stopwords*, a Figura 7 apresenta a nuvem de palavras do acervo da revista TEIAS com todos os termos que ocorreram duas ou mais vezes.

REDES DE COLABORAÇÃO

A colaboração entre pesquisadores pode ocorrer das mais diferentes maneiras, porém, elas somente são identificáveis por metadados nas coautorias e nas citações (referências) relatadas nos trabalhos. No que diz respeito a este trabalho, a análise de redes de colaboração foi realizada através da construção de grafos com base nas coautorias dos trabalhos (OLIVEIRA, 2012; NEWMAN, 2004a; NEWMAN, 2004b; SEMAAN *et al.*, 2019a).

Koehler *et al.* (2015) aborda estudos de comunidades, tema de grande aderência no contexto de análises de redes de colaboração. Em uma visão mais ampla, antes do advento das tecnologias modernas de comunicação, a noção de centralidade em redes já era utilizada para análises sociais em empresas (BAVELAS, 1950). De fato, tais abordagens não são recentes, e são complexas devido ao caráter dinâmico e suas diferentes configurações no tempo e no espaço.

Conforme Bavelas (1950), “[...] *num grupo de pessoas, um indivíduo que se encontra estrategicamente localizado num caminho mais curto de comunicação entre pares de indivíduos, está numa posição mais central da rede [...]*”. Assim, tal indivíduo é considerado influente devido à sua localização privilegiada, sendo ele “[...] *responsável por transmitir, modificar, ou reter a informação entre membros do grupo [...]*”.

Identificação de pesquisadores considerados influentes

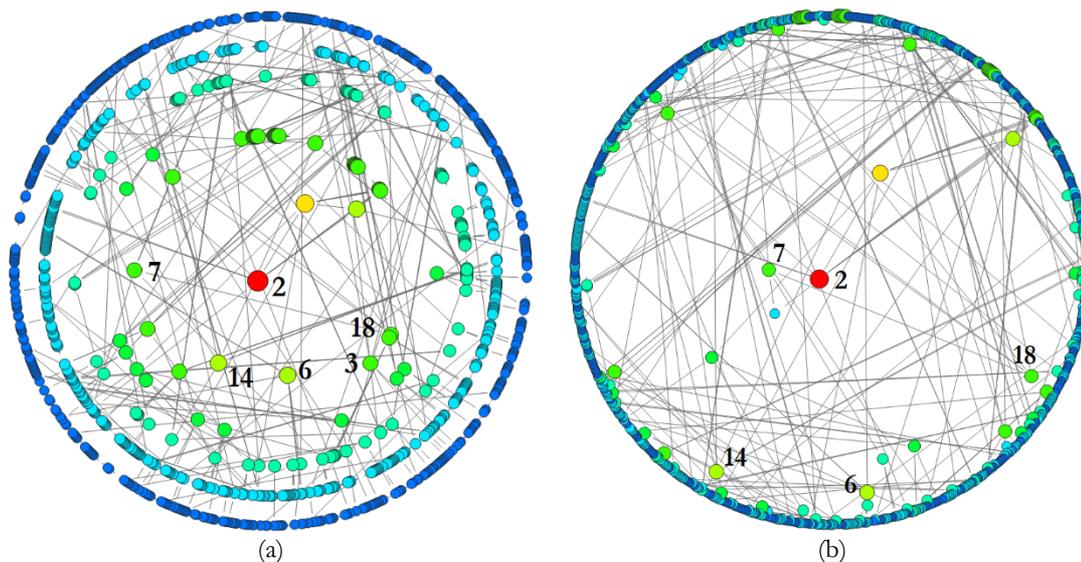
As estatísticas gerais, embora de grande importância em relação a meta-análises, não são suficientes para fornecer uma visão holística. Assim sendo, de forma a possibilitar diferentes perspectivas em estudos relacionados a comunidades, várias medidas de centralidades em redes podem ser consideradas. Na presente pesquisa foram consideradas as centralidades de grau e de intermediação, e busca-se mensurar a influência dos autores em uma comunidade conforme sua posição estrutural no grafo (BORGATTI; EVERETT, 2006; SEMAAN *et al.*, 2019a; BAVELAS, 1950; BEAUCHAMP, 1965).

A Figura 8 apresenta o Grafo 1 representado em dois *layouts* diferentes: (a) centralidade de grau e (b) centralidade de intermediação. O Grafo foi construído com todos os 940 autores (vértices) após o processo de unificação e consolidação de seus nomes, além das relações de colaboração entre pares, que compõem 803 arestas. Em ambas as representações gráficas, tanto as cores quanto os tamanhos dos vértices estão relacionadas aos seus graus. No *layout* radial quando mais próximo do centro um vértice está localizado, mais importante (influyente) ele é, conforme o critério considerado. Os números existentes nessa figura correspondem aos identificadores (coluna *Id*) dos autores reportados na Tabela 1, de modo a destacá-los como influentes na comunidade da revista TEIAS. De maneira adicional, ressalta-se que os grafos não contemplaram a frequência de publicação dos autores.

A Tabela 1 apresenta os 20 autores considerados mais influentes no acervo, conforme os critérios determinados no presente estudo. A seleção ocorreu pelo critério maior frequência de publicação (coluna *Freq*) e em seguida, foram relacionados com a coluna Grau, que indica a quantidade de coautores que participaram desses trabalhos e a coluna CI, que indica a posição do autor conforme o cálculo da medida de centralidade de intermediação. As colocações indicadas na coluna CI estão em um contexto em que existem 940 autores, sem contar os autores publicaram trabalhos somente de maneira isolada (SA), não contemplados nos grafos.

Ainda em relação à Tabela 1, os nomes dos pesquisadores estão em ordem alfabética, uma vez que o trabalho não tem a intenção de avaliar os pesquisadores, embora tais resultados prestem uma singela homenagem pela contribuição à comunidade. Ainda nesse sentido, de modo adicional, a Figura 9 apresenta os nomes de todos os autores que possuem, ao menos, dois trabalhos publicados, em um total de 781 autores.

Figura 8: Grafos em *layout* radial por centralidade de grau (a) e de intermediação (b).



Fonte: Autoria própria

Figura 9: Autores com ao menos dois artigos publicados nos acervos analisados.



Fonte: Autoria própria

Tabela 1: autores considerados mais influentes na comunidade da Revista TEIAS.

Id	Nome	Freq.	Grau	CI
1	Adriana Hoffmann Fernandes	5	3	55 ^a
2	Antonio Amorim	6	9	1 ^a
3	Carlos Roberto Carvalho	9	5	13 ^a
4	Clara Pereira Coutinho	4	2	90 ^a
5	Fabiany Tavares Silva	5	3	14 ^a
6	Geovana Mendonça Lunardi	6	6	9 ^a
7	Janete Magalhães Carvalho	7	5	2 ^a

8	Jean Houssaye	4	1	91 ^a
9	João Batista Bottentuit Junior	4	4	43 ^a
10	Laura Noemi Chalh	4	3	20 ^a
11	Maria Luiza Sussekind	5	2	69 ^a
12	Maria Teresa Esteban	6	3	17 ^a
13	Marisol Barenco Mello	5	2	35 ^a
14	Nilda Alves	6	6	12 ^a
15	Paula Almeida de Castro	4	2	46 ^a
16	Rafael Marques Gonçalves	4	3	23 ^a
17	Rita Ribes	7	5	25 ^a
18	Sandra Kretli Silva	5	5	10 ^a
19	Stela Caputo	5	3	92 ^a
20	Virgínia de Oliveira Silva	4	1	93 ^a

Fonte: autoria própria

CONCLUSÕES

O presente trabalho teve como o objetivo principal fornecer perspectivas sobre a revista TEIAS por meio de meta-análises com base em seu acervo. Uma vez que foram considerados metadados coletados do portal da revista, destaca-se a grande importância de repositórios científicos serem abertos à comunidade com simplicidade de acesso, organização e transparência. Ressalta-se, portanto, o empenho de toda a equipe da revista que, além de possuir trabalhos de alta qualidade e relevância, comprovada pela avaliação Qualis da CAPES, contribuem efetivamente com a democratização do conhecimento.

O acervo da revista, no momento da coleta dos dados, possuía 56 edições em 21 anos de publicações, com 1.016 artigos publicados por 1.323 autores após a consolidação dos dados. A seção Metodologia apresentou o processo de KDD bem como conceitos introdutórios de Teoria dos Grafos, considerados nas Análises de redes de colaboração.

Na seção Revisão da literatura e perspectivas foram relatados trabalhos relevantes que abordam repositórios de importantes revistas e eventos. Nela também estão as estatísticas gerais da revista, em que é possível observar o crescimento dos quantitativos de autores e de artigos ao passar do tempo. De maneira adicional, o gráfico crescimento do acervo ilustra o total de trabalhos acumulados ao longo dos anos.

Mesmo antes da apresentação dos grafos que ilustram as colaborações entre autores, fortes indícios do aumento da colaboração entre eles podem ser observados ainda com base nas estatísticas gerais. A média de autores por artigo aumentou de cerca de 1 em 2000 para 2 em 2018. Além disso, o percentual de autores isolados, que chegou a 100% dos artigos publicados em 2004, foi reduzido para cerca de 25% dos trabalhos publicados em 2018. Adicionalmente, cerca de 84% dos autores possuem apenas um trabalho no acervo e 53,8% publicaram artigos de maneira isolada. Entretanto, cerca de 46% os trabalhos possuem entre dois e quatro coautores.

Em relação aos termos presentes nos títulos das publicações, além da apresentação das 20 palavras mais frequentes por ano, os termos *Educação*, *Formação* e *Professor* possuíram as maiores médias de ocorrência por ano. De modo adicional, o termo *Escolar* ocorreu no título de trabalhos publicados em 19 dos 21 anos da revista. Uma nuvem de palavras apresenta todos os termos dos títulos, com exceção das *stopwords* e das palavras com apenas uma ocorrência.

Nas análises de redes de colaborações foram apresentados grafos com *layout* radial por centralidade de grau e de intermediação. A Tabela 1 relaciona os autores considerados mais influentes conforme os critérios adotados. Adicionalmente, uma nuvem de palavras contém os nomes dos pesquisadores que publicaram mais de uma vez na revista, e fornece uma significativa (simbólica) imagem da comunidade da Revista TEIAS.

As perspectivas apresentadas contemplaram a apresentação de estatísticas gerais, análises em relação aos termos frequentes em títulos dos trabalhos, as colaborações através das redes modeladas em grafos e indicaram, inclusive, autores considerados influentes. A abordagem considerou análises objetivas e sem intencionalidade política, restritas aos dados coletados em seu portal, e o trabalho apresenta-se como uma importante referência para pesquisas futuras.

Dentre os novos caminhos a serem investigados, destacam-se: análises das palavras em destaque *versus* o contexto temporal; análise de grupos de uma mesma instituição e as colaborações interinstitucionais; uso das referências (citações) como fonte de origem para a construção de grafos e cálculo de medidas de centralidade; o uso de outras modelagens com grafos (ponderados ou dígrafos); e, por fim, contemplar também bases de dados de outros eventos importantes no cenário nacional.

REFERÊNCIAS

- ALVES, Márcio Fagundes; FLORENTINO, Rita de Cássia; DIAS, Rosilana Aparecida. Informática educativa e protagonismo juvenil: o projeto “Jovens navegando pela cidade”. *Revista TELAS*, v. 8, n. 14-15, 2007.
- ARANHA, Christian Nunes. *Uma abordagem de pré-processamento automático para mineração de textos em português: sob o enfoque da inteligência computacional*. Tese. (Doutorado), Pontifícia Universidade Católica do Rio de Janeiro, Rio de Janeiro, 2007.
- BAVELAS, Alex. Communication patterns in task-oriented groups. *Journal of the Acoustical Society of America*, n. 22, 1950.
- BEAUCHAMP, Murray A., *An improved index of centrality*, Behavioral Science, v. 10, 1965.
- BOAVENTURA NETTO, Paulo Oswaldo. *Grafos: teoria, modelos, algoritmos*. Editora Blucher, 2001.
- BORGATTI, Stephen P.; EVERETT, Martin G. *A Graph-theoretic perspective on centrality*. Social Networks. 28 (4), 2006.
- BOTTENTUIT Jr, João Batista; LISBÔA, Eliana Santana; COUTINHO, Clara Pereira. Narrativas digitais na formação inicial de professores: um estudo com alunos de licenciatura em pedagogia. *Revista TELAS*, v. 13, n. 27, 2012.
- CAVALCANTI, Thiago Rodrigues; DA SILVA, Fábio Q. B. *Historical, conceptual, and methodological aspects of the publications of the brazilian symposium on software engineering: a systematic mapping study*. *Anais do 25th Brazilian Symposium on Software Engineering (SBES)*, São Paulo, 2011.
- KOEHLER, Cristiane; MACHADO-SPENCE, Nádie Christina; CARVALHO, Marie Jane Soares. A constituição de grupos, agrupamentos e comunidades em websites de redes sociais. *Revista TELAS*, v. 16, n. 43, 2015.
- FERNANDES, Clovis Torres; SANTOS, Neide. Pesquisa e desenvolvimento em informática na educação no brasil – parte i. *Revista Brasileira de Informática na Educação*, v. 4, 1999.

- FREEMAN, Linton C. A set of measures of centrality based on betweenness. *Sociometry*, v. 40, n. 1, p. 35-41, 1977.
- FREEMAN, Linton C. Centrality in social networks: conceptual clarification. *Social Networks*, 1978.
- GOLDSCHMIDT, Ronaldo; PASSOS, Emmanuel Lopes. *Data Mining: um guia prático*. Rio de Janeiro: Elsevier. 2005.
- LIMA, Lucas Henrique Costa *et al.* The collaboration network of the Brazilian Symposium on Databases - 30 editions of history. *Journal of the Brazilian Computer Society* 23:10, 2017.
- MAGALHÃES, Cleyton V. C *et al.* Caracterizando a Pesquisa em Informática na Educação no Brasil: Um Mapeamento Sistemático das Publicações do SBIE. XXIV Simpósio Brasileiro de Informática na Educação - SBIE, 2013.
- NEWMAN, Mark E. J. *Coauthorship networks and patterns of scientific collaboration*. In: *Proceedings of the National Academy of Sciences*, 2004a.
- NEWMAN, Mark E. J. *The structure of scientific collaboration networks*. Proc. Natl. Acad. Sci. USA, v. 98, 2001.
- NEWMAN, Mark E. J. *Who is the best connected scientist? a study of scientific coauthorship networks*. In: BEN-NAIM, Eli; FRAUENFELDER, Hans; TOROCZKAI, Zoltan. (ed.). *Complex Networks*. Springer Berlin / HEIDELBERG, 2004, (Lecture Notes in Physics, v. 650), 2004b.
- OLIVEIRA, Edvaldo Artmann. *Sobre a colaboração na comunidade de sistemas de informação através dos simpósios SBSI*. Dissertação. (Mestrado), UNIRIO, Rio de Janeiro, 2012.
- PROCACI, Thiago B *et al.* *Estudo exploratório das produções e colaborações entre pesquisadores em informática na educação: uma análise de publicações do simpósio brasileiro de informática na educação de 2001 a 2013*. SBIE, 2015.
- RODRIGUES, Natan S.; RALHA, Célia G. Conhecendo a comunidade de sistemas de informação no Brasil: um estudo comparativo utilizando diferentes abordagens de banco de dados. XI *Brazilian Symposium on Information System*, Goiânia-GO, 2015.
- SEMAAN, Gustavo Silva *et al.* Um panorama geral das contribuições e das redes de colaboração do SBIE. *Simpósio Brasileiro de Informática na Educação - SBIE*, 2019(a).
- SEMAAN, Gustavo Silva *et al.* Um panorama dos trabalhos do SIMPEP: 25 anos de contribuições. *Simpósio de Engenharia de Produção - SIMPEP*, 2019(b).
- SEMAAN, Gustavo Silva *et al.* Uma análise das redes de colaboração do SIMPEP. *Simpósio de Engenharia de Produção - SIMPEP*, 2019(c).
- TEIAS, 2020, Portal da Revista TEIAS, Disponível em <https://www.e-publicacoes.uerj.br/index.php/revistateias>. Acesso em 6 mar. 2020.
- VANZ, Samile Andrea de Souza. *As redes de colaboração científica no Brasil (2004-2006)*. Tese. (Doutorado), Universidade Federal do Rio Grande do Sul, Porto Alegre, 2009.

*Submetido em julho de 2020
Aprovado em fevereiro de 2021*

Informações dos autores

Gustavo Silva Semaan
Universidade Federal Fluminense (UFF)
E-mail: gustavosemaan@id.uff.br
ORCID: <https://orcid.org/0000-0003-2873-2628>
Link Lattes: <http://lattes.cnpq.br/4519888592231795>

Edelberto Franco Silva
Universidade Federal de Juiz de Fora (UFJF)
E-mail: edelberto@ice.ufjf.br
ORCID: <https://orcid.org/0000-0002-0058-9260>
Link Lattes: <http://lattes.cnpq.br/3987091765361506>

José André de Moura Brito
Escola Nacional de Ciências Estatísticas do Instituto Brasileiro de Geografia e Estatística (ENCE-IBGE)
E-mail: jambrito@gmail.com
ORCID: <https://orcid.org/0000-0002-2825-0058>
Link Lattes: <http://lattes.cnpq.br/9036541085964477>