

Metodologia para análise de redes de coautoria a partir de fontes de informação que não contemplam metadados padronizados

Methodology for analyzing co-authorship networks based on information sources that do not include standardized metadata

Marcos de Souza ¹mds@ufba.br

Submetido em: 03-02-2024

Aceito em: 05-11-2024

Resumo: A pesquisa apresenta uma metodologia de Análise de Rede Social (ARS) para Rede de Coautoria de um *corpus* de documentos científicos sem metadados padronizados e formatos específicos para *softwares* de ARS, como os disponíveis em bases de dados. Um dos objetivos específicos inclui a Análise de Conteúdo dos documentos. O *corpus*, com 44 documentos publicados entre 2007 e 2023, foi coletado a partir do currículo na Plataforma Lattes do autor, organizado em diretórios e importado para o *software Mendeley*, onde metadados dos documentos foram preenchidos e corrigidos. Posteriormente, os metadados foram exportados em formato .bib ou .ris para importação, configuração e visualização da Rede de Coautoria usando o *software VosViewer*. A Rede de Coautoria apresentou 26 itens entre autores e coautores, organizados em oito *clusters*, com 57 links e 118 links fortes. Entre as relações mais fortes, estão Marcos de Souza e Fernanda Gomes Almeida do *Cluster 1*, com 9 pesquisas, destacando temas como Modelagem de Tópicos, Arquitetura da Informação e Gestão de Dados de Pesquisa. O estudo demonstra possibilidades para ARS a partir de outras fontes de informação, além dos metadados fornecidos por bases de dados, ampliando o leque para pesquisas específicas, como identificar a rede de coautoria em diferentes áreas do conhecimento, departamentos, cursos de graduação ou programas de pós-graduação, líderes de grupos de pesquisa, grupos de trabalho de eventos ou em toda uma universidade, entendendo as relações entre *clusters*.

Palavras-chave: Análise de redes sociais; rede de coautoria; *Corpus*; *Mendeley*; *Vosviewer*.

¹ Professor Adjunto do Instituto de Ciência, Tecnologia e Inovação (ICTI), da Universidade Federal da Bahia (UFBA), na área de Humanidades Digitais e Métodos Computacionais da Pesquisa Social. Doutorado em Gestão e Organização do Conhecimento pela Universidade Federal de Minas Gerais (UFMG), com foco em Gestão Tecnológica. Mestrado em Cognição e Linguagem pela Universidade Estadual Norte Fluminense (UENF), com ênfase em Pesquisas Interdisciplinares em Comunicação, Educação e Novas Tecnologias da Informação. Especializações em Informática na Educação pelo Instituto Federal do Espírito Santo (IFES), em Docência do Ensino Superior pelo Centro Universitário São Camilo - Espírito Santo (CeUSC), e em Desenvolvimento de Aplicações para WEB pelo Centro de Ensino Superior de Juiz de Fora (CESJF); Graduação em Sistemas de Informação pelo CeUSC.

Abstract: The research aims to present a Social Network Analysis (SNA) methodology for the Co-authorship Network of a corpus of scientific documents without standardized metadata and specific formats for SNA software, such as those available in databases. One of the specific objectives includes content analysis of the documents. The corpus, with 44 documents published between 2007 and 2023, was collected from the author's Lattes Platform CV, organized into directories and imported into the Mendeley software, where document metadata was filled in and corrected. Subsequently, the metadata was exported in .bib or .ris format for import, configuration and visualization of the Co-authorship Network using the VosViewer software. The Co-authorship Network showed 26 items between authors and co-authors, organized into eight clusters, with 57 links and 118 strong links. Among the strongest relationships are Marcos de Souza and Fernanda Gomes Almeida from Cluster 1, with 9 research projects, highlighting topics such as Topic Modeling, Information Architecture and Research Data Management. The study demonstrates possibilities for ARS from other sources of information, in addition to the metadata provided by databases, broadening the range for specific research, such as identifying the co-authorship network in different areas of knowledge, departments, undergraduate courses or postgraduate programs, leaders of research groups, working groups of events or in an entire university, understanding the relationships between clusters.

Keywords: Social network analysis; coauthorship network; *corpus*; *Mendeley*; *VosViewer*.

1 INTRODUÇÃO

A Análise de Redes Sociais (ARS) refere-se a uma abordagem de pesquisa que tem ganhado popularidade nos últimos anos em diferentes áreas do conhecimento (Recuero, 2017), sendo abordada, como exemplo, por pesquisadores da Administração, Antropologia, Biologia, Comunicação, Matemática e Medicina (Cunha; Amaral; Dantas, 2015).

Trata-se de uma técnica de coleta e análise de dados relacionados às redes sociais, possibilitando, assim, uma análise sistemática a partir de sua estrutura e de medidas específicas. Além disso, seu uso permite obter representações das redes sociais a partir de grafos, que são constituídos por nós - atores conectados por ligações - relações entre os componentes (Cunha; Amaral; Dantas, 2015).

Com raízes na Sociometria e na Teoria dos Grafos, os estudos de diferentes estruturas de ARS passaram a receber notoriedade com o aumento da disponibilidade de dados de interações sociais que passaram a ser analisados por softwares de mídia social (Recuero; Bastos; Zago, 2015).

Dentro do contexto da comunicação científica, existem serviços em bases de dados como a *Web of Science*, *Scopus*, *Dimensions*, *Lens* e *PubMed* que disponibilizam conjuntos padronizados de metadados de pesquisa em formatos específicos para serem analisados por *softwares* de ARS.

Todavia, outras publicações como anais de eventos, pesquisas publicadas em periódicos e livros, bem como teses e dissertações não indexadas em bases de dados que disponibilizam conjuntos de metadados padronizados e em formatos específicos para ARS também são fontes de informação que podem contribuir diretamente para o fazer da ciência. A saber, as bases de dados exigem listas de critérios rigorosos para a indexação de periódicos científicos.

Partindo deste princípio, o problema da pesquisa se apresenta como: de que maneira seria possível conduzir uma ARS de pesquisas científicas publicadas na internet, que não disponibilizam conjuntos de metadados padronizados e formatos específicos, a exemplo dos metadados disponibilizados em base de dados, para a utilização em software de ARS?

Essa questão se apresenta como um desafio relevante, independente da área do conhecimento, uma vez que pesquisas acadêmicas não indexadas em bases de dados como *Web of Science*, *Scopus*, *Dimensions*, *Lens* e *PubMed* e que não disponibilizam conjuntos de metadados padronizados e em formatos específicos para ARS também desempenham um papel crucial na disseminação do conhecimento acadêmico e podem preencher lacunas para a compreensão das interações científicas em uma escala mais ampla.

A pesquisa tem como objetivo geral apresentar uma metodologia para ARS de um *corpus* de documentos científicos que não apresentam conjuntos de metadados padronizados e em formatos específicos para *softwares* de ARS, como os que são disponibilizados em bases de dados. Alguns exemplos são as pesquisas disponibilizadas no *Google Scholar*, na Base de Dados em Ciência da Informação (Brapci) ou mesmo nos periódicos científicos que utilizam o *Open Journal Systems* (OJS). Dentre os objetivos específicos, destacam-se: a) demonstrar as etapas de preenchimento, correção e normalização dos metadados dos documentos analisados; b) detalhar as etapas para a construção de uma Rede de Coautoria, utilizando os metadados do *corpus* de documentos e o *software* de ARS *VosViewer*; c) apresentar e discutir *clusters* como resultados alcançados por meio da ARS do *corpus* estudado.

Acredita-se, como pressuposto, que estudos de ARS em pesquisas que não são indexadas em bases de dados ou que não possuem mecanismos de exportação

de metadados padronizados também podem contribuir em diferentes áreas do conhecimento. Por exemplo, por meio de publicações em anais de eventos, é possível identificar as redes de coautoria, acoplamento bibliográfico ou cocitação que mais contribuem para uma determinada área do conhecimento. Da mesma forma, em publicações de um determinado Programa de Pós-graduação, é possível identificar a distância média entre os colaboradores, autores que mais colaboram, grupos que colaboram com mais intensidade e o nível de colaboração da comunidade envolvida. Além disso, em publicações de periódicos, é possível identificar as temáticas que estão em evidência por meio da coocorrência de termos.

Justifica-se a importância da pesquisa uma vez que a contribuição da ARS aos estudos cientométricos, enquanto abordagem, tem ganhado espaço na comunicação científica, especialmente em pesquisas descritivas, que podem ser realizadas durante a fase que antecede a qualificação em programas *lato* e *stricto sensu* de Pós-Graduação. Com isso, é possível realizar o levantamento bibliográfico e identificar, por meio de indicadores de ARS, as redes de coautoria que mais contribuem para uma determinada área do conhecimento. A partir daí, estabelece-se um direcionamento para o desenvolvimento do referencial teórico ou revisão da literatura a ser utilizado nas respectivas pesquisas.

A ARS também pode ser utilizada por pesquisadores com outros objetivos, como, por exemplo, descrever as publicações de um determinado grupo de pesquisa, buscando compreender suas conexões, bem como as relações entre docentes e discentes, docentes e docentes, discentes e discentes de Programas de Pós-graduação. Além disso, pode ser empregada para descrever a produção de um único pesquisador ou escritor.

2 REFERENCIAL TEÓRICO

O referencial teórico, dividido em duas subseções, engloba os conceitos, características e software utilizado na ARS, bem como conceitos relativos às Redes de Coautoria.

2.1 Análise de redes sociais

Dentre as perspectivas de estudo de grupos sociais está a ARS, que possibilita a realização de uma análise sistemática a partir de sua estrutura e medidas específicas (Recuero; Bastos; Zago, 2015). Essa abordagem tem origens nas Ciências Sociais (Freeman, 1996; Marteleto, 2001), Sociologia, Sociometria e na Teoria dos Grafos, sendo apoiada por um viés matemático utilizado para analisar as relações sociais (Kozinets, 2015; Recuero; Bastos; Zago, 2015).

A ARS começou a ser utilizada no início do século XX. Posteriormente a sua difusão na sociologia passou a ser aplicada em áreas como Antropologia, Biologia, Comunicação, Matemática e Medicina (Cunha; Amaral; Dantas, 2015). A análise de redes estabelece um novo paradigma na pesquisa sobre a estrutura social (Marteleto, 2001), buscando identificar as relações complexas entre vários elementos de um sistema social em diferentes dimensões.

A ARS pode ser representada por matrizes - uma construção matemática dos dados da rede social - ou por grafos - uma representação visual da rede (Recuero, 2017). Essas representações revelam as estruturas sociais e os sujeitos que a integram (Cunha; Amaral; Dantas, 2015; Freeman, 1996; Marteleto, 2001).

Em uma representação visual da estrutura social de um determinado grupo de agentes, podem ser consideradas posições específicas, grau de conectividade e os recursos em circulação, dependendo do fluxo de comunicação (Freeman, 1996; Marteleto, 2001).

A técnica de coleta e análise de dados relacionados às redes sociais tem possibilitado a representação dessas redes em diagramas, compostos por nós - que representam os atores da rede - e ligações - que representam as relações entre os componentes, sejam eles pessoas físicas ou jurídicas (Cunha; Amaral; Dantas, 2015; Kozinets, 2015).

Existem diferentes categorias de redes sociais, como as redes formais e informais apontadas por Lazzarini, Chaddad e Neves (2000) e Azevedo e Rodriguez (2010), além das redes sociotécnicas, implícitas e abertas destacadas por Pisciotta (2006). Outros tipos de redes sociais incluem as internas, externas e hierárquicas, também apresentadas por Azevedo e Rodriguez (2010).

A premissa dos estudos das estruturas sociais diz respeito aos atores sociais inseridos em estruturas complexas e suas relações uns com os outros. A ARS pressupõe a percepção de um grupo social em rede, o que possibilita a análise por meio de premissas metodológicas (Recuero, 2017).

Para estudar os comportamentos e opiniões dos atores sociais, faz-se necessário a existência das estruturas nas quais os indivíduos se inserem. Entretanto, a unidade de análise não são os atributos individuais, como sexo, idade ou gênero, mas sim o conjunto de relações que os atores estabelecem por meio de suas interações (Marteleto, 2001).

A ideia das estruturas sociais e seus elementos, enquanto papel fundamental no comportamento e na visão do mundo dos indivíduos, é destacada por Recuero (2017, p. 9):

Cada indivíduo, desse modo, está inserido em grupos sociais que são constituídos a partir de suas relações nos mais variados ambientes (por exemplo, família, escola, amigos ou trabalho). Essas relações são estabelecidas por interações e associações e vão conferir aos atores determinadas posições nas suas redes sociais, que vão sendo modificadas por essas mesmas ações. A posição desses atores é, ao mesmo tempo, produto e produtora de interações, ou seja, a rede influencia e é influenciada pela posição de seus usuários (Recuero, 2017, p. 9).

A ARS tem ganhado relevância na comunidade científica por analisar relacionamentos entre entidades sociais e identificar os padrões e implicações desses relacionamentos (Azevedo; Rodriguez, 2010). Cunha, Amaral e Dantas (2015) destacam os tipos mais utilizados de representações por grafos na ARS: rede não orientada - não apresentam ligações com direções previamente definidas; rede orientada - ligações definidas; rede ponderada - com pesos, forças e parâmetros de orientação definidos; rede multigrafo - nós contendo várias ligações entre eles; rede de autointerações - interações entre o próprio nó; e rede de grafo completo - nós ligados uns aos outros sem autointeração.

Dentre as principais características estruturais de uma rede, encontram-se: ator - representando cada indivíduo, grupos de indivíduos, corporações, setores ou departamentos interligados à rede; ligações - representações gráficas de linhas que conectam os atores; subgrupos - subconjuntos de atores de uma rede,

independentemente do tamanho; relação - um tipo específico de elo de ligação entre membros de um determinado grupo; tamanho - a quantidade de conexões existentes entre os atores de uma rede; densidade - o quociente entre o número de ligações existentes e o número de ligações possíveis em uma rede; distância geodésica - a menor distância entre dois atores (Azevedo; Rodriguez, 2010; Lago Júnior, 2005; Souza; Quandt, 2008).

Dependendo do escopo estudado, as redes sociais podem ser constituídas por milhares de atores conectados por diferentes vínculos e situações. Para um estudo de redes sociais de forma ampla, abrangendo toda a sua estrutura, faz-se necessário o uso de softwares específicos que auxiliem no mapeamento, controle e análise dos dados (Azevedo; Rodriguez, 2010). A ARS tem se destacado pela utilização de tecnologias computacionais e da internet, pois permite diferentes maneiras de gerar representações estruturais das redes sociais (Limas; Dantas; 2009).

Azevedo e Rodriguez (2010) destacam que o uso de softwares de ARS requer conhecimentos específicos por parte do pesquisador, como conhecimento em ARS, compreensão das necessidades do estudo ou organização, conhecimento do tipo de rede a ser estudada, bem como familiaridade com as principais características e recursos dos softwares.

Silva *et al.* (2021) ressaltam que os softwares de ARS oferecem ao pesquisador uma ampla variedade de possibilidades para compreender as dinâmicas das relações entre atores de uma rede, permitindo análises quantitativas ou integração com outros métodos, por meio da abordagem qualitativa.

A maioria dos softwares de ARS possibilita o armazenamento dos dados, que posteriormente podem ser descritos e manipulados por meio de matrizes, permitindo a visualização dos grafos. Além disso, no mesmo contexto, eles permitem a modificação da cor e da forma dos nós das redes de acordo com as características específicas de cada grupo, possibilitando também a visualização e a análise cruzada de vários dados da rede, fornecendo uma análise visual e matemática das relações informais (Azevedo; Rodriguez, 2010).

Dentre os softwares de ARS, tanto *open source* quanto licenciados, destacam-se: *Cfinder*, *CiteSpace*, *GraphViz*, *Guess*, *InFlow*, *JUNG*, *MultiNet*,

NetDraw, NetMiner, NodeXL, Pajek, SocNetV, UCINET, VosViewer. Orta (2016) apresenta uma lista com mais de cem ferramentas, incluindo aplicações, softwares, pacotes e *addons* (complementos), juntamente com suas respectivas descrições para ARS.

2.2 Colaboração científica e redes de coautoria

No contexto acadêmico, a estrutura dinâmica de colaboração em Redes de Coautoria tem ganhado notoriedade e despertado o interesse de instituições avaliadoras e de fomento em ciência e tecnologia (Balancieri *et al.*, 2005). Esse interesse é particularmente evidente em áreas como a Ciência da Informação, que aborda estudos no campo da bibliometria e da cientometria, contribuindo para essa tendência (Liu *et al.*, 2005).

A colaboração entre pesquisadores no desenvolvimento de pesquisas tornou-se uma das características distintivas da ciência moderna. Um dos indicadores mais utilizados para avaliar a colaboração científica é o de coautoria (Katz; Martin, 1997). Os autores também destacam que a colaboração científica pode se manifestar de diferentes maneiras, formalizadas ou informalizadas, como na coorientação de trabalhos acadêmicos ou na escrita conjunta de projetos.

A coautoria tem sido amplamente utilizada como medida de colaboração científica entre pesquisadores ou instituições (Oliveira; Grácio, 2008), na qual dois ou mais atores contribuem para a criação de uma produção científica (Matheus; Vanz; Moura, 2007).

Autores como Cronin (2005), Katz e Martin (1997) e Newman (2004) destacam a coautoria como um indicador de suma importância para a rede de colaboração, especialmente em publicações do canal formal da comunicação científica, como artigos em periódicos científicos. Por outro lado, Vanz e Stumpf (2010) apontam que as redes de coautoria representam uma faceta da colaboração científica, uma vez que nem todas as colaborações entre coautores resultam em publicações.

No entanto, a análise de redes de colaboração tem sido utilizada para compreender os vínculos entre atores, instituições e países, pois possibilita a

verificação dos dados por outros atores e a exploração de grandes amostras (Katz; Martin, 1997).

O fato é que as análises de coautorias acadêmicas oferecem uma visão sobre a estrutura e dinâmica das relações com base nas colaborações entre os pesquisadores, sendo representadas por grafos (Mena-Chalco; Digiampietri; Cesar-Júnior, 2012). Bufrem, Gabriel Júnior e Sorribas (2011, p. 7) destacam que "uma rede de coautoria é uma rede na qual os nós são os pesquisadores e há conexão entre eles sempre que compartilham a autoria de um artigo".

Nesse contexto de rede, os resultados das análises de coautoria são de suma importância para o entendimento dos diversos aspectos relacionados à colaboração. Eles podem, por exemplo, identificar a colaboração de toda uma comunidade de pesquisa, grupos de pesquisadores que trabalham de forma isolada ou autores que mais colaboram com uma determinada área ou grupo. Tais informações podem contribuir para a tomada de decisões em um contexto científico (Bordin; Gonçalves; Tedesco, 2014).

Estudos em fontes de informações especializadas são fundamentais para aumentar a visibilidade das redes de coautoria. Dessa forma, os atores podem identificar as redes nas quais colaboram, os laços mais fortes e os elos mais fracos, entre outros elementos que constituem as redes sociais (Silva, 2014). Oliveira e Grácio (2008) destacam que os estudos sobre redes de colaboração são relevantes "[...] na medida em que proporcionam visibilidade à produção científica, à análise do seu domínio e aos cientistas mais produtivos, entre outros objetivos" (Oliveira; Grácio, 2008, p. 36).

3 METODOLOGIA

Para a construção do referencial teórico, essencial para o desenvolvimento da pesquisa, foram utilizados artigos científicos e livros disponibilizados por meio do Portal de Periódicos da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), *Scielo* e *Google Scholar*.

O *corpus* de documentos utilizado para ARS foi construído em duas etapas. A primeira consistiu na extração das informações das publicações no currículo Lattes

do pesquisador utilizado como referência nesta pesquisa. A segunda etapa envolveu a busca dos documentos publicados na internet pelo autor em periódicos científicos, repositórios e anais de eventos, seguida do *download* dos respectivos documentos, que foram organizados em um diretório.

Dentre os documentos coletados estão resumos, resumos expandidos e artigos completos publicados em anais de eventos, além de resumos e artigos completos publicados em periódicos científicos, bem como capítulos de livros avaliados por comitês científicos. Foram excluídas pesquisas monográficas como livros, trabalhos de conclusão de curso de graduação e especializações *lato* e *stricto sensu*. A constituição do *corpus* de documentos ocorreu entre os dias 9 e 15 de maio de 2023.

O tratamento dos metadados dos documentos que constituem o *corpus* foi realizado por meio do software gerenciador de referências denominado *Mendeley*². Essa etapa envolveu o preenchimento, a correção e a normalização dos metadados dos documentos. Para a ARS, foi utilizado o software *VosViewer*³, responsável pela construção e visualização das redes bibliométricas utilizadas neste estudo. Diferente da rede de coautoria total, que apresenta a rede completa entre todos os autores conectados, a rede de coautoria resultante foi do tipo egocêntrico, sob o ponto de vista de um indivíduo central. Já a análise da rede considerou critérios específicos de limites populacionais (Garton; Haythornthwaite; Wellman, 1997).

Para a análise dos resultados e discussões, foi adotada a abordagem de Análise de Conteúdo, apresentada por Bardin (2015) e Cunha, Amaral e Dantas (2015) e concentrando-se a análise em dados de identificação de autoria e coautoria, títulos, resumos e palavras-chave dos documentos utilizados, bem como informações complementares contidas na plataforma Lattes dos coautores. Nesta etapa, não foram utilizados quaisquer softwares para criar categorias. Em vez disso, foram aplicadas abordagens qualitativas e quantitativas com base nos documentos analisados, a partir dos grafos gerados.

² Mendeley – software gerenciador de referências. Disponível em: <https://www.mendeley.com/>. Acesso em: 09 jun. 2023.

³ VosViewer - software para construção e visualização de redes bibliométricas. Disponível em: <https://www.vosviewer.com/>. Acesso em: 09 jun. 2023.

A definição do autor e de suas pesquisas utilizadas para construção do *corpus* de documentos ocorre exemplificativamente, uma vez que o objetivo da pesquisa é destacar as etapas e processos para a ARS de um *corpus* de documentos científicos que não apresentam conjuntos de metadados padronizados e em formatos específicos para softwares de ARS.

Entretanto, a escolha do autor utilizado e suas respectivas pesquisas é justificada com base em sua formação interdisciplinar e na rede de coautorias estabelecidas em diferentes áreas do conhecimento ao longo do período analisado, tornando-se possível observar nos resultados, de maneira visual, as redes sociais e as respectivas relações. Além disso, cabe citar a familiaridade com o percurso acadêmico analisado e a relação direta do autor com os coautores das pesquisas do *corpus* de documentos, justificando assim a escolha da análise por conveniência, o que não deslegitima a pesquisa.

4 RESULTADOS E DISCUSSÕES

Os resultados e discussões são apresentados em duas subseções. A primeira descreve as etapas necessárias para a geração da rede de coautoria e a segunda subseção apresenta os resultados, análises e discussões da rede de coautoria.

4.1 Da coleta e manipulação de dados, apresentação de resultados e discussões iniciais

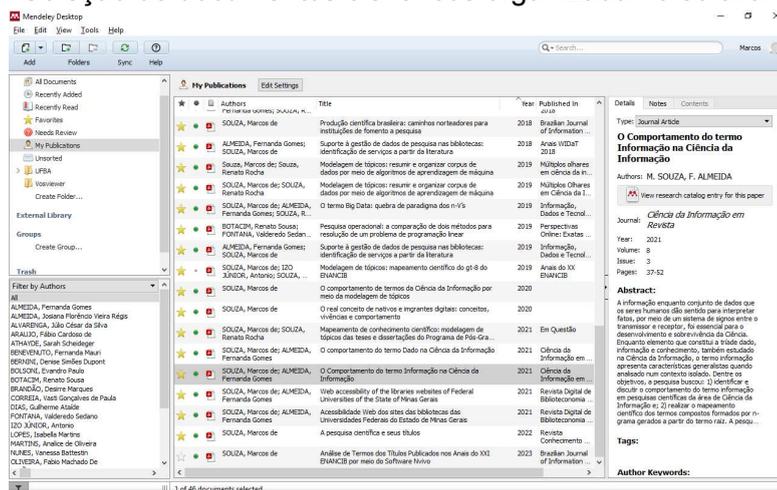
O *corpus* de documentos⁴ utilizado como exemplo para a realização da análise de redes de coautoria consiste em um total de 44 pesquisas. O período das publicações analisadas foi de 2007 a 2023, correspondendo ao ano da primeira e da última publicação realizada pelo autor escolhido até o momento.

Os documentos que compõem o *corpus* foram importados para o software gerenciador de referências *Mendeley*, onde foi possível realizar o

⁴ *Corpus* de documento. Lista contendo os dados dos documentos textuais utilizados nesta pesquisa para a Análise de Redes de Coautoria, juntamente com o link de acesso aos respectivos documentos. Disponível em: <https://doi.org/10.6084/m9.figshare.23114714.v1>. Acesso em: 09 jun. 2023.

preenchimento/correção dos metadados de cada arquivo. Na Figura 1, na parte central, o documento intitulado “O Comportamento do Termo Informação na Ciência da Informação”, enquanto à direita são exibidos os metadados que foram preenchidos/corrigidos do respectivo artigo.

Figura 1 – Coleção de documentos científicos organizada no software Mendeley



Fonte: Elaborado pelo autor (2023).

Dentre os principais campos apresentados pelo software estão: tipo de documento (livro, capítulo de livro, estudo de caso, programa de computador, documentos de conferências, enciclopédias, filmes, áudios, artigos de periódicos, artigos de revistas, artigos de jornais, patente, relatório, estatuto, programas de televisão, teses, páginas da web, documentos de trabalho); título; autores; periódico; ano; volume; número; páginas; resumo; palavras-chave; URI; DOI; ISSN ou ISBN.

A etapa de preenchimento/correção dos metadados é de suma importância para a normalização dos dados e, conseqüentemente, para apresentar resultados mais coesos. Esta etapa inclui a normalização dos nomes de autores e coautores, uma vez que "SOUZA, Marcos de" é diferente de "SOUZA, Marcos" ou "SOUZA, M.". Sem a normalização dos metadados, o mesmo autor ou coautor utilizado como exemplo pode aparecer nos resultados duas ou três vezes, dividindo o peso de suas publicações e apresentando resultados divergentes.

Após a normalização dos metadados do *corpus*, foi necessário exportar o documento contendo o conjunto de metadados⁵ da coleção de documentos, que pode ser nos formatos *BibTex*⁶ (.bib) ou RIS⁷ – *Research Information Systems* (.ris). Esse documento foi posteriormente importado pelo software *VosViewer*. Para realizar essa etapa, foi necessário selecionar os documentos que compõem a coleção e clicar no menu "*File > Export...*".

Para a construção da rede de coautoria, realizada por meio do *software VosViewer* e utilizando os metadados da coleção de documentos organizada pelo *software Mendeley*, foram realizadas as seguintes etapas:

a) Clicar no botão "*Create...*" e escolher o tipo de dados a ser importado. Marcar a opção "*Create a map based on bibliographic data*" com a descrição "*Choose this option to create a co-authorship, keyword co-occurrence, citation, bibliographic coupling, or co-citation map based on bibliographic data*".

b) Posteriormente, é necessário marcar a opção "*Read data from reference manager files*". A descrição do campo informa os tipos de formatos suportados, que incluem RIS, *EndNote* e *RefWorks*.

c) Em seguida, localizar o arquivo com a extensão .bib ou .ris, contendo os metadados dos documentos exportados pelo *software* gerenciador de referência, que será importado para o *VosViewer*.

d) A seguir, foi definido em quatro etapas o tipo de análise e método de contagem "*Choose type of analysis and counting method*". O primeiro campo refere-se à definição do tipo de análise, que no caso deste estudo é a análise de coautoria. O segundo campo refere-se ao método de contagem, que também é usado nesta pesquisa, sendo a contagem por totalidade. O terceiro campo é opcional e está relacionado a um dicionário de sinônimos, denominado pelo *software* como "*Thesaurus file*". Neste caso, utiliza-se um arquivo do tipo .txt, que contém as informações para normalizar os dados, caso essa normalização não tenha sido

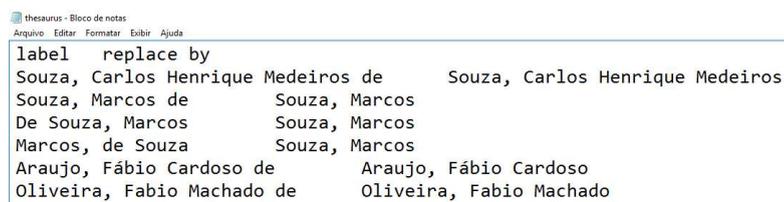
⁵ Metadados no formato .RIS utilizado para a construção da ARS. Disponível em: <https://doi.org/10.6084/m9.figshare.23114714.v1>. Acesso em: 09 jun. 2023.

⁶ BibTex - uma ferramenta e um formato de arquivo que são usados para descrever e processar listas de referências. Disponível em: <http://www.bibtex.org/>. Acesso em: 09 jun. 2023.

⁷ RIS - é um formato de tag padronizado desenvolvido para permitir que programas de citação troquem dados. Disponível em: [https://en.wikipedia.org/wiki/RIS_\(file_format\)](https://en.wikipedia.org/wiki/RIS_(file_format)). Acesso em: 09 jun. 2023.

realizada no software de gerenciamento de referências durante a etapa de preenchimento/correção. A Figura 2 apresenta um exemplo de dicionário de sinônimos.

Figura 2 – Exemplo de dicionário de sinônimos



label	replace by
Souza, Carlos Henrique Medeiros de	Souza, Carlos Henrique Medeiros
Souza, Marcos de	Souza, Marcos
De Souza, Marcos	Souza, Marcos
Marcos, de Souza	Souza, Marcos
Araujo, Fábio Cardoso de	Araujo, Fábio Cardoso
Oliveira, Fabio Machado de	Oliveira, Fabio Machado

Fonte: Elaborado pelo autor (2023).

O quarto campo desta etapa também é opcional e está relacionado a ignorar ou não os documentos com mais de 25 autores, permitindo a personalização desse limite quantitativo.

e) A tela de configuração seguinte apresenta o número mínimo de documentos por autor. Por exemplo, do total de 44 documentos que constituem o *corpus*, existem 26 itens – autores e coautores. Desse conjunto de coautores, nove deles participaram de duas pesquisas, um participou de cinco pesquisas, dois deles participaram de sete pesquisas, um deles participou de oito pesquisas, um deles participou de nove pesquisas, e um deles participou de dez pesquisas, conforme apresentado na Figura 3, que também apresenta a lista dos autores, podendo ser ordenados pela ordem alfabética, pelo número de documentos em que o autor e coautor participam ou pela força total do *link*. Neste momento, é possível realizar novamente a conferência dos nomes dos autores. Caso exista alguma divergência, poderá ser corrigida no dicionário de sinônimos.

Figura 3 – Lista de autores que compõem o *corpus* de documentos



Selected	Author	Documents	Total link strength
<input checked="" type="checkbox"/>	souza, marcos de	44	74
<input checked="" type="checkbox"/>	brandão, desire marques	8	22
<input checked="" type="checkbox"/>	botacim, renato souza	7	17
<input checked="" type="checkbox"/>	souza, carlos henrique medeiros de	7	14
<input checked="" type="checkbox"/>	souza, renato rocha	10	14
<input checked="" type="checkbox"/>	athayde, sarah scheideger	5	12
<input checked="" type="checkbox"/>	almeida, fernanda gomes	9	11
<input checked="" type="checkbox"/>	oliveira, fabio machado de	2	8
<input checked="" type="checkbox"/>	soares, elíelton sopeleto	2	7
<input checked="" type="checkbox"/>	toffano, José Carlos dias	2	7
<input checked="" type="checkbox"/>	xavier, bruno missi	2	7
<input checked="" type="checkbox"/>	martins, analice de oliveira	2	5
<input checked="" type="checkbox"/>	almeida, josiana florêncio vieira régis	2	4
<input checked="" type="checkbox"/>	bermini, denise simões dupont	2	4
<input checked="" type="checkbox"/>	correia, vasti gongalves de paula	2	4
<input checked="" type="checkbox"/>	izo júnior, antonio	2	4
<input checked="" type="checkbox"/>	benevenuto, fernanda maui	1	3
<input checked="" type="checkbox"/>	fontana, valderedo sedano	1	3
<input checked="" type="checkbox"/>	lopes, isabella martins	1	3
<input checked="" type="checkbox"/>	alvarenga, ilídio cesar da silva	1	2

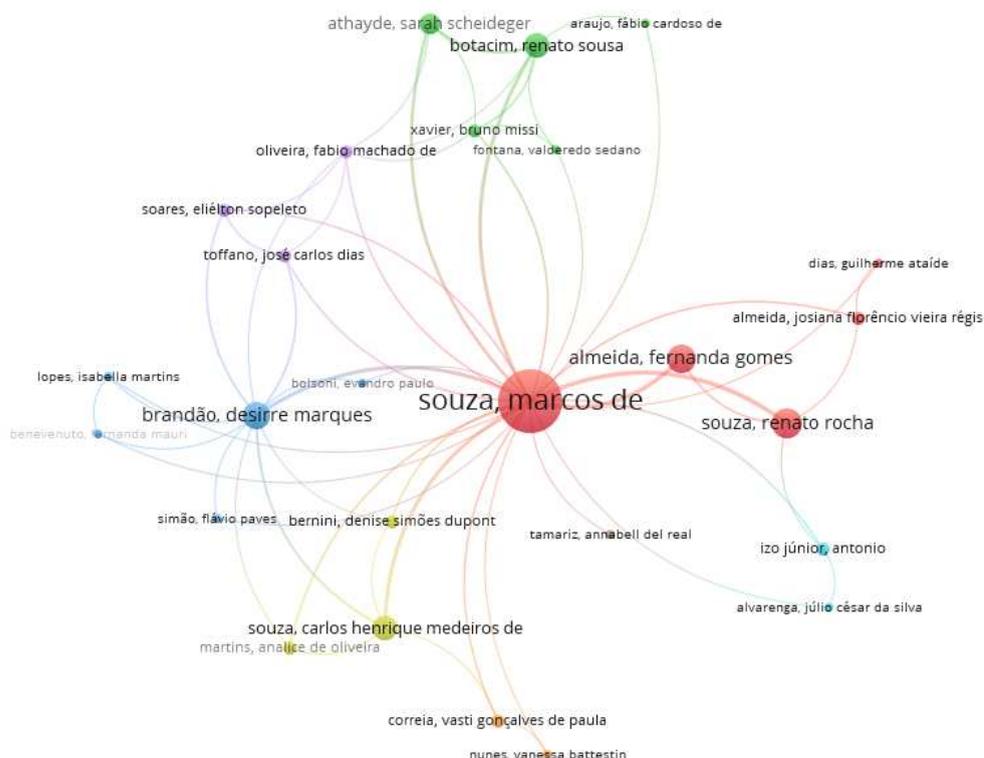
Fonte: Elaborado pelo autor (2023).

f) A última etapa está relacionada à escolha do número de autores e seus respectivos *links* de coautoria. Quanto menor for o número de autores selecionados, menor será a rede de coautoria, e vice-versa. Para cada um dos 26 autores, será calculada a força total dos vínculos de coautoria com outros autores. Os autores com a maior força total de link serão selecionados. Para exibir toda a rede, basta adicionar o valor um no campo número de autores selecionados.

Após clicar em finalizar, a rede de coautoria é gerada com base nos dados exportados do software de gerenciamento de referências e importados para o *VosViewer*, conforme apresentado na Figura 4. Cabe ressaltar que a opção de ajuste dos resultados pode ser realizada por meio do dicionário de sinônimos.

Para entendimento, os *clusters* são agrupados com base nos itens - autores e links - que possuem força, indicando o número de coautorias dos autores, no caso da análise em questão. Os itens podem possuir diferentes atributos de peso e pontuação, representados por valores não negativos. O peso de um item pode indicar a sua importância dentro da rede. Quanto maior o peso, maior a relevância do item para o *cluster* e maior destaque do elemento na visualização da rede. Todos os agrupamentos e cálculos são realizados pelo *VosViewer* (Van Eck; Waltman, 2023).

Figura 4 – Visualização da Rede de coautoria

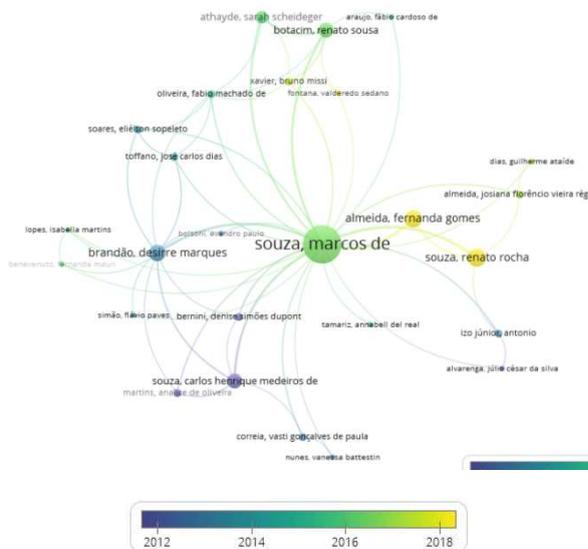


Fonte: Elaborado pelo autor (2023).

A rede de coautoria gerada pelo software *VosViewer* contemplou 26 itens, 57 links, força total de 118 links e foi organizada em oito *clusters* de agrupamentos por coautorias. Os links e a força total dos links são considerados pelo software como dois atributos de peso padrão. Para um determinado item, esses atributos representam, respectivamente, o número de vínculos de coautoria e a força total dos links de coautoria. Já os *clusters*, são agrupamentos de conjuntos de itens (Van Eck; Waltman, 2023).

O software *VosViewer* possibilita também outros dois tipos de visualizações, sendo o primeiro denominado *Overlay Visualization*, que apresenta os *clusters* separados por períodos, com as parcerias mais antigas representadas pela cor lilás e as parcerias mais recentes representadas pela cor amarela, conforme apresentado na Figura 5.

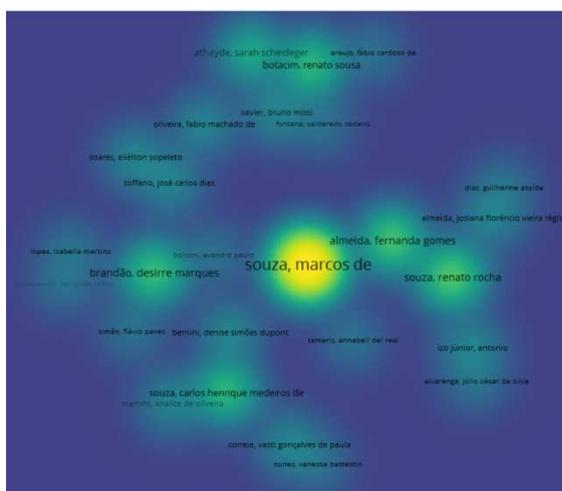
Figura 5 – Visualização da Rede de coautoria por períodos de publicação



Fonte: Elaborado pelo autor (2023).

A visualização do tipo *Density Visualization* exibe a densidade dos dados. Quanto maior o número de itens na vizinhança de um ponto e quanto maiores os pesos desses itens, mais amarela será a cor do ponto. Por outro lado, quanto menor o número de itens na vizinhança de um ponto e quanto menores os pesos dos desses itens, mais azul será a cor do ponto, conforme apresentado na Figura 6.

Figura 6 – Visualização da Rede de coautoria por densidade



Fonte: Elaborado pelo autor (2023).

Os gráficos dinâmicos⁸ da rede de coautoria podem ser acessados e explorados por meio de navegadores Web, utilizando diferentes recursos, como *zoom in* e *zoom out*, informações de cada autor como itens, links e força total de links, *clusters*, atualização de *layout*, alteração das cores para o padrão publicado por ano, inversão da cor de fundo para indivíduos com fotossensibilidade, aumento da espessura das linhas, captura de imagem da tela, compartilhamento, entre inúmeras possibilidades e combinações de recursos.

A combinação dos três tipos de visualização da rede - *cluster*, tempo e densidade - amplia as possibilidades de análises detalhadas do *corpus* de documentos estudado. É possível, por exemplo, analisar a identificação das relações da rede e seu grau de contribuição em um determinado período.

4.2 Da análise de conteúdo dos resultados

Em relação à análise de conteúdo da produção científica que compõe o *corpus* de documentos, o *Cluster 1*, destacado na Figura 4, tornou-se evidente ao considerar os atributos de maior peso e pontuação, mediante o volume de publicações entre os anos de 2017 a 2023, bem como a correlação entre autores e coautores. O *Cluster* em questão pode ser dividido em dois momentos distintos, mas correlacionados. O primeiro está relacionado às pesquisas publicadas entre 2017 a 2020, anos do doutoramento do item central do *Cluster 1* no Programa de Pós-graduação em Gestão e Organização do Conhecimento (PPG-GOC), realizado na Universidade Federal de Minas Gerais (UFMG). O segundo momento é destacado entre os anos de 2021 a 2023, após a realização do doutoramento.

As pesquisas relacionadas à primeira fase foram construídas em parcerias com os seguintes pesquisadores: ALMEIDA, F. G.; ALMEIDA, J. F. V. R.; DIAS, G.A.; SOUZA, R.R (orientador). Os assuntos pesquisados e publicados nos formatos de artigos completos em periódicos, artigos completos e resumos expandidos em anais de eventos e capítulos de livros estão relacionados às disciplinas cursadas pelo item central do *Cluster 1* durante a sua passagem pelo PPG-GOC, tanto na linha de

⁸ Gráfico dinâmico do resultado de pesquisa – rede de coautoria. Disponível em: <https://bit.ly/3BPmTyn>. Acesso em: 09 jun. 2023.

pesquisa em Gestão & Tecnologia da Informação e Comunicação⁹ (GETIC) quanto na linha de pesquisa em Arquitetura e Organização do Conhecimento (AOC).

Tal correlação ocorre entre disciplinas cursadas, como Estudos Avançados em Informação e Conhecimento, Tópicos Especiais em Arquitetura e Organização do Conhecimento, Representação Ontológica em Sistemas Formais, Sistemas de Recuperação da Informação, Tópicos Especiais em Gestão e Tecnologia I e II, e as publicações realizadas durante o período de doutoramento, a saber: DIAS, G. A.; ALMEIDA, J. F. V. R.; SOUZA, M. "O uso do termo ontologia em comunicações orais no gt-2 do ENANCIB" (2017); SOUZA, M.; ALMEIDA, J. F. V. R.; SOUZA, R. R. "Ontologias para interfaces de sistemas" (2017); SOUZA, M.; SOUZA, R. R. "A arquitetura e a recuperação da informação para a melhoria do índice h na plataforma Lattes" (2017); SOUZA, M.; SOUZA, R. R. "Fluxo informacional para redução de dados duplicados na plataforma Lattes" (2017); SOUZA, M.; SOUZA, R. R. "Produção científica brasileira e as instituições de fomento à pesquisa" (2017); SOUZA, M. "Produção científica brasileira: caminhos norteadores para instituições de fomento à pesquisa" (2018); ALMEIDA, F. G.; SOUZA, M. "Suporte à gestão de dados de pesquisa nas bibliotecas: identificação de serviços a partir da literatura" (2018 e 2019); SOUZA, M.; ALMEIDA, F. G.; SOUZA, R. R. "O termo Big Data: quebra de paradigma dos n-Vs" (2018 e 2019).

Dentre os itens com maiores destaques do *Cluster 1* estão: SOUZA, R. R. com 10 documentos, quatro links e 14 links fortes; e ALMEIDA, F. G. com nove documentos, dois links e 11 links fortes.

Já as pesquisas publicadas relacionadas à temática abordada na tese defendida no ano de 2020 pelo pesquisador central do *Cluster 1* - Souza, Marcos de, intitulada "O comportamento de termos da Ciência da Informação por meio de Modelagem de Tópicos," estão: SOUZA, M.; SOUZA, R. R. "Modelagem de tópicos: resumir e organizar *corpus* de dados por meio de algoritmos de aprendizagem de máquina" (2019); SOUZA, M.; IZO JUNIOR, A.; SOUZA, R. R. "Modelagem de tópicos: mapeamento científico do gt-8 do ENANCIB" (2019); SOUZA, M.; SOUZA, R. R. "Mapeamento de conhecimento científico: modelagem de tópicos das teses e

⁹ Descrição das linhas de pesquisas do PPG-GOC: Disponível em: <https://ppggoc.eci.ufmg.br/area-de-concentracao-e-linhas-de-pesquisa/>. Acesso em: 09 jun. 2023.

dissertações do Programa de Pós-graduação em Ciência da Informação da UFMG" (2020 - data do aceite), 2021 (data da publicação).

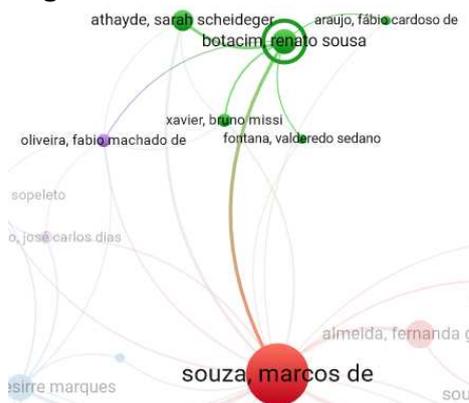
O segundo momento a ser destacado no *Cluster* 1 diz respeito às pesquisas publicadas após o doutoramento, sendo três delas publicações originadas também da tese de doutorado: SOUZA, M.; ALMEIDA, F. G. "O comportamento do termo Dado na Ciência da Informação" (2021); SOUZA, M.; ALMEIDA, F. G. "O comportamento do termo Informação na Ciência da Informação" (2021); SOUZA, M.; ALMEIDA, F. G. "O comportamento do termo Conhecimento na Ciência da Informação" (2023).

Os autores Marcos de Souza e Fernanda Gomes Almeida ainda desenvolveram outra pesquisa que uniu a experiência e a formação de suas respectivas áreas de atuação, sendo Sistemas de Informação e Biblioteconomia: SOUZA, M.; ALMEIDA, F. G. "Acessibilidade Web dos sites das bibliotecas das Universidades Federais do Estado de Minas Gerais" (2021). Os dois pesquisadores desenvolveram em parceria um total de nove pesquisas.

A partir deste momento, SOUZA, M. desenvolveu pesquisas na área da escrita científica e análise qualitativa, destacando as publicações em periódicos científicos: SOUZA, M. "A pesquisa científica e seus títulos" (2022); SOUZA, M. "Análise de termos dos títulos publicados nos anais do XXI ENANCIB por meio do software NVivo" (2023).

O *Cluster* 2 – rótulo verde, tem como item de maior força o pesquisador Renato Sousa Botacim, com sete documentos, seis links e 17 links fortes, conforme apresentado na Figura 7. Os elementos centrais do *Cluster* 2 – Botacim, Renato Sousa e *Cluster* 1 – Souza, Marcos de, publicaram pesquisas, incluindo artigos completos publicados em periódicos, trabalhos completos publicados em anais de congressos, resumos expandidos publicados em anais de congressos e apresentações de trabalhos em eventos, durante os anos de 2015 a 2019.

Figura 7 – Cluster 2 – rótulo verde



Fonte: Elaborado pelo autor (2023).

Referente ao período dos documentos que constituem o *Cluster 2*, Marcos de Souza atuava como professor e coordenador dos cursos de Análise e Desenvolvimento de Sistemas (ADS) e Sistemas de Informação (SI) no Centro Universitário São Camilo – Espírito Santo (CeUSC). As pesquisas desenvolvidas por BOTACIM, R.S. e SOUZA, M., em parceria com os demais coautores, também estão relacionadas às disciplinas ministradas pelo docente, incluindo Tópicos Especiais em Tecnologia da Informação, Computador e Sociedade, Gestão de Projetos e Interface Homem Máquina.

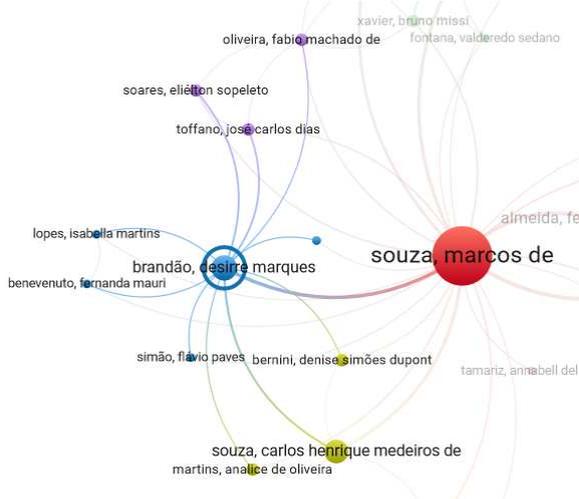
Além disso, é possível observar uma continuidade de estudos voltados ao comportamento dos nativos e imigrantes digitais (apresentado no *Cluster 3*) em pesquisas realizadas em parceria com os discentes do CeUSC. Dentre as publicações estão: SOUZA, M.; BOTACIM, R. S.; ARAUJO, F. C. “Gerenciamento de projetos e gestão de pessoas com diversidades culturais no jogo Clash of Clans” (2015); BOTACIM, R. S.; ATHAYDE, S. S.; OLIVEIRA, F. M.; XAVIER, B. M.; SOUZA, M. “Computação nas nuvens: evolução e peculiaridade dos serviços e da segurança da informação” (2016); BOTACIM, R.S.; ATHAYDE, S. S.; SOUZA, M. “As tecnologias da informação e comunicação como instrumento de ensino e aprendizado: o precipício entre imigrantes e nativos digitais” (2016); BOTACIM, R. S.; ATHAYDE, S. S.; SOUZA, M. “Infográficos: a combinação das Leis da Gestalt no processo criação e desenvolvimento” (2016); BOTACIM, R. S.; ATHAYDE, S. S.; SOUZA, M. “A proposta de uma aplicação com infraestrutura em grid: o uso de

sistemas distribuídos no compartilhamento de recursos computacionais” (2016); ATHAYDE, S. S.; BOTACIM, R. S.; SOUZA, M. “O uso das TICs nos processos de ensino e aprendizado: as diferenças entre nativos e imigrantes digitais” (2016).

Além dessas pesquisas, no ano de 2019, foi publicado um artigo completo em periódico científico, entretanto, fora da temática de pesquisa do docente, a saber: BOTACIM, R. S.; FONTANA, V. S.; XAVIER, B. M.; SOUZA, M. “Pesquisa Operacional: a comparação de dois métodos para resolução de um problema de programação linear” (2019).

O *Cluster 3* – rótulo azul, possui como item mais forte a pesquisadora Desirre Marques Brandão, com oito documentos, 11 links e 22 links fortes, conforme apresentado na Figura 8. Os elementos centrais do *Cluster 3* – Brandão, Desirre Marques e *Cluster 1* - Souza, Marcos de, possuem publicações dos tipos artigos completos em periódicos, capítulos de livros, trabalhos completos publicados em anais de eventos e resumos expandidos publicados em anais de eventos realizados entre os anos de 2011 a 2016, destacando as fases de pré-mestrado, mestrado e pós-mestrado de SOUZA, M.

Figura 8 – Cluster 3 – rótulo azul



Fonte: Elaborado pelo autor (2023).

O primeiro momento refere-se à publicação de uma pesquisa realizada durante o período do curso de Pós-graduação *lato sensu* em Docência do Ensino

Superior¹⁰ do CeUSC, que antecede o período de mestrado, destacando a publicação: SOUZA, M.; BERNINI, D. S. D.; BRANDAO, D. M.; SOUZA, C. H. M. “Modalidades avaliativas e seus objetivos - Análise das ferramentas EAD utilizadas na disciplina Metodologia do Trabalho Científico do Centro Universitário São Camilo Espírito Santo” (2019).

Durante o período da Pós-graduação em Cognição e Linguagem (PGCL), nível mestrado, realizado pela Universidade Estadual do Norte Fluminense (UENF), SOUZA, M.; BRANDAO, D. M. desenvolveram pesquisas em coautoria com outros pesquisadores, a saber: SOUZA, M.; MARTINS, A. O. (orientadora de Brandão); SOUZA, C. H. M. (orientador de Souza); BRANDAO, D. M. “Desenvolvimento colaborativo de textos jornalísticos utilizando o serviço web Google Drive: produção, leitura e circulação de textos” (2012); SOUZA, M.; BRANDAO, D. M.; BOLSONI, E. P. (*in memoriam*) “Elaboração de artigo científico quantitativo: uma abordagem sobre a coleta, tabulação e estatística de dados e serviço web google drive spreadsheet” (2013); SOUZA, M.; BRANDAO, D. M.; SOUZA, C. H. M. “Web Google Drive Spreadsheet: abordagem sobre coleta, tabulação e estatística de dados para pesquisa científica utilizando métodos de pesquisa quantitativa” (2013); BRANDAO, D. M.; SIMAO, F. P.; SOUZA, M. “Impressões dos alunos do curso técnico em informática sobre a leitura e a produção textual utilizando Tecnologias de Informação e Comunicação” (2014).

Por se tratar de um *cluster* que apresenta o segundo maior número de links fortes no escopo estudado, e, conseqüentemente, um número elevado de coautores quando comparado aos demais *clusters*, torna-se possível identificar diferentes conexões do *Cluster 3* com outros conjuntos gerados pelo VosViewer com base nos dados de pesquisa, destacando o *Cluster 4* e o *Cluster 5*.

Durante o período de realização do mestrado, estão as publicações de SOUZA, M. com coautores apresentados no *Cluster 4* – rótulo amarelo, sendo: SOUZA, C. H. M.; SOUZA, M. “A utilização das ferramentas tecnológicas como instrumento facilitador de processos nas avaliações internas das instituições de

¹⁰ Descrição do curso de Pós-Graduação *lato sensu* em Docência do Ensino Superior do CeUSC. Disponível em: <http://www.saocamilo-es.br/poscachoieiro/educacao/docencia-do-ensino-superior.html/>. Acesso em: 09 jun. 2023.

ensino superior” (2011); SOUZA, M.; SOUZA, C. H. M. “FeedBurner como ferramenta de controle e disseminação da informação nas redes sociais virtuais das instituições do ensino superior” (2012); SOUZA, M.; CORREIA, V. G. P.; SOUZA, C. H. M. “O real nativo e imigrante digital nas redes sociais digitais” (2013).

As publicações realizadas durante este período estão alinhadas com a linha de Pesquisas Interdisciplinares em Comunicação, Educação e Novas Tecnologias da Informação¹¹ (PICENTI) do PGCL, e, conseqüentemente, com as disciplinas cursadas por SOUZA, M., entre os anos de 2011 e 2013, tais como: Tópicos Especiais em Educação Inclusiva e Acessibilidade; Estatística e Pesquisa em Educação; Pesquisa Discente; Linguagem e Novas Tecnologias da Comunicação; Introdução a Cognição e Linguagem; O Texto e as Novas Tecnologias – Modos de Produção, Leitura e Circulação; Produção Científica Interdisciplinar.

As publicações em periódicos científicos após a obtenção do título de mestre em Cognição e Linguagem pela UENF, realizadas por SOUZA, M. com diferentes coautores, estão apresentadas tanto no próprio *Cluster 3* – rótulo azul quanto no *Cluster 5* – rótulo lilás. As temáticas das pesquisas dão continuidade aos estudos realizados durante o período de mestrado sobre o comportamento dos nativos e imigrantes digitais nas redes sociais digitais, associadas ou alinhadas com as disciplinas ministradas por SOUZA, M. para os cursos de graduação em ADS e SI do CeUSC, incluindo Computador e Sociedade, Tópicos Especiais em Tecnologia da Informação e Gestão de Projetos.

Dentre as pesquisas destacam-se: BRANDÃO, D. M.; SOARES, E. S.; TOFFANO, J. C. D.; SOUZA, M. “Vlogs monetizados: do hobby ao profissionalismo utilizando as redes sociais digitais” (2014); BRANDAO, D. M.; SOARES, E. S.; TOFFANO, J. C. D.; SOUZA, M.; OLIVEIRA, F. M. “Do hobby a atividade profissional nas redes sociais digitais: uma abordagem aos vlogs de entretenimento;” (2014); SOUZA, M.; BRANDAO, D. M.; LOPES, I. M.; BENEVENUTO, F. M. “Gestão de projetos e gerenciamento de riscos no jogo clash of clans” (2016).

O *Cluster 6* – rótulo azul piscina, possui como item mais forte o pesquisador Antonio Izo Júnior, com dois documentos, três links e quatro links fortes, conforme

¹¹ Descrição da linha de pesquisa PICENTI do PGCL. Disponível em: <http://pgcl.uenf.br/o-programa#apresentacao-e-linhas-de-pesquisa/>. Acesso em: 09 jun. 2023.

apresentado na Figura 9. Os elementos centrais do *Cluster 6* – Izo Júnior, Antonio e *Cluster 1* – Souza, Marcos de, possuem publicações dos tipos capítulo de livro e trabalhos completos publicados em anais de eventos, respectivamente em 2007 e 2019.

Figura 9 – *Cluster 6* – rótulo azul piscina



Fonte: Elaborado pelo autor (2023).

Dentre as publicações, destacam-se: SOUZA, M.; ALVARENGA, J. C. S. (*in memoriam*); IZO JUNIOR, A. “A utilização do processo de arquitetura da informação na melhoria da qualidade dos portais web das IES” (2007); SOUZA, M.; IZO JUNIOR, A.; SOUZA, R. R. “Modelagem de tópicos: mapeamento científico do gt-8 do ENANCIB” (2019).

A publicação referente ao ano de 2007 foi resultado de um trabalho de orientação realizado por ALVARENGA, J. C. S. durante a primeira Pós-graduação *lato sensu* em Desenvolvimento de Sistemas para Web cursada por SOUZA, M. entre os anos de 2005 e 2006 junto ao Centro de Ensino Superior de Juiz de Fora (CESJF). Na ocasião, foi estudada a temática da Arquitetura da Informação da área da Ciência da Informação, área esta em que SOUZA, M. realizaria o doutorado entre os anos de 2017 a 2020. Já a publicação de 2019 retoma a parceria entre IZO JUNIOR, A. e SOUZA, M., dessa vez com as orientações de SOUZA, R. R. (orientador de SOUZA, M. no doutorado), justificando assim a interface entre os *Clusters 6* e 1.

Torna-se importante ressaltar que diversas outras análises podem ser realizadas com base nos resultados apresentados pela ARS, dependendo dos objetivos da pesquisa. Por exemplo, é possível explorar os *clusters* com

documentos, links e links de força com menor valor, ou examinar as relações entre as coautorias, além do item principal de cada cluster.

Clusters com valores menos expressivos podem conter informações relevantes para uma pesquisa, como o *Cluster 7* – rótulo laranja, que possui como item de maior relevância a pesquisadora Vasti Gonçalves de Paula Correia, orientadora de Marcos de Souza durante a realização do curso de Pós-graduação *lato sensu* em Informática na Educação¹² ofertado pelo Instituto Federal do Espírito Santo (IFES). Além disso, os autores colaboraram em um capítulo de livro elaborado pelos membros centrais do *Cluster 7* e do *Cluster 1*, referenciado com a publicação: SOUZA, M.; CORREIA, V. G. P.; NUNES, V. B. “O real nativo e imigrante digital nas redes sociais digitais” (2013).

Os autores também estabeleceram parcerias com o pesquisador Carlos Henrique Medeiros de Souza, item principal do *Cluster 4*, gerando a publicação de um artigo científico em periódico, ultrapassando os limites físicos entre instituições de ensino e, principalmente, a relação entre as áreas de Educação, Tecnologia, Cognição e Linguagem.

A Rede de Coautoria apresentada oferece uma visão do grafo sobre a estrutura e dinâmica das relações com base nas colaborações entre os pesquisadores (Mena-Chalco; Digiampietri; Cesar-Júnior, 2012) destacando os nós as conexões que compartilham a autoria de documentos (Bufrem; Gabriel Júnior; Sorribas, 2011) para o entendimento dos diversos aspectos relacionados à colaboração (Bordin; Gonçalves; Tedesco, 2014).

5 CONSIDERAÇÕES FINAIS

O problema da pesquisa foi respondido ao apresentar as etapas e processos necessários para conduzir uma ARS de pesquisas científicas disponibilizadas na internet, que não possuem mecanismos para exportação de conjuntos de metadados nas plataformas onde estão hospedadas. Isso ocorre de maneira

¹² Descrição do curso de Pós-Graduação *lato sensu* em Informática na Educação do IFES. Disponível em:

<https://cef.ifes.edu.br/index.php/cursos/167-cursos/pos-graduacao/16308-pos-graduacao-em-informatica-na-educacao/>. Acesso em: 09 jun. 2023.

diferente em bases de dados que permitem que essas informações sejam disponibilizadas em formatos padronizados e aceitos por softwares de ARS, como o VosViewer.

Quanto ao objetivo geral e específicos “a” e “b”, que foram atingidos nesta pesquisa, cabe mencionar as etapas de preenchimento e correção de metadados dos documentos, bem como a exportação dos metadados do *corpus* de documentos analisados, utilizando o software gerenciador de referências Mendeley. Além disso, o processo inclui a importação desses metadados, a configuração e a correção dos dados, se necessário, utilizando tesouros, até a visualização da rede, incluindo a possibilidade de exploração por meio de uma versão interativa.

Diferente das bases de dados, que apresentam conjuntos de metadados padronizados e possibilitam a ARS em escala maior e/ou generalista, a análise sugerida nesta pesquisa utiliza um *corpus* de documentos específicos, com metadados organizados utilizando software de gerenciamento de referências e a elaboração de rede de coautoria por meio de *software* de ARS. Essa abordagem permite realizar pesquisas em diferentes amplitudes, dependendo do *corpus*, mas com características específicas, como identificar a rede de coautoria em diferentes áreas do conhecimento, departamentos, cursos de graduação ou programas de pós-graduação, líderes de grupos de pesquisa, grupos de trabalho de eventos ou mesmo em toda uma universidade para entender as relações entre *clusters*.

Além disso, é possível realizar a ARS em documentos de diferentes fontes de informação, que também contribuem para o fazer da ciência, como, por exemplo, os anais de um evento científico com avaliação por pares, que disponibilizam os documentos em um ambiente que não permite a exportação dos metadados dos documentos para realização de estudos de ARS.

No que tange ao objetivo “c”, foi realizada a análise de conteúdo do *corpus* de documentos, entre os anos de 2007 e 2023, utilizado para a ARS, especificamente na Rede de Coautoria, explorando elementos como título, autoria e coautoria, resumos e palavras-chave dos documentos que compõem o *corpus*, dentre outras informações disponibilizadas nos currículos dos coautores das pesquisas na Plataforma Lattes.

Além disso, foram apresentadas as características dos *clusters*, destacando o item central de cada *cluster*, o número de documentos, links e links fortes de cada grupo, suas relações, e geradas as respectivas discussões. Também foi possível identificar diferentes assuntos abordados pelo pesquisador central da coleção em coautoria com outros pesquisadores, tais como: Arquitetura da Informação, Fluxo Informacional, Ontologias, Produção Científica, Gestão de Dados de Pesquisa, Comportamento de Usuários, Gestão de Projetos, Acessibilidade Web, Modelagem de Tópicos, Leitura e Produção Textual e Tecnologias da Informação e Comunicação.

O pressuposto da pesquisa foi comprovado, uma vez que foi utilizado um conjunto de documentos científicos de diferentes fontes de informação que não estavam indexadas em bases de dados que fornecem mecanismos de exportação de metadados dos documentos para estudos de ARS.

Especificamente para o conjunto de documentos em questão, tornou-se possível identificar uma diversidade de assuntos mediante a análise de conteúdo e à formação interdisciplinar do autor, que possui graduação em Sistemas de Informação, mestrado em Cognição e Linguagem, doutorado em Gestão e Organização do Conhecimento, além de três pós-graduações *lato sensu* nas áreas de educação e tecnologias.

Para pesquisas futuras, sugere-se a realização de outros tipos de ARS nos documentos que compõem o *corpus*, como acoplamento bibliográfico, cocitação de referências e coocorrência de termos. Além disso, como extensão do estudo, seria interessante expandir a rede, inserindo informações sobre pesquisas realizadas pelos coautores do *corpus* estudado em colaboração com suas respectivas redes, a fim de entender melhor o contexto dos pesquisadores.

REFERÊNCIAS

AZEVEDO, T. B.; RODRIGUEZ, M. V. R. Softwares para Análise de Redes Sociais - ARS. *In*: CONGRESSO NACIONAL DE EXCELÊNCIA EM GESTÃO, 6., 2010, Niterói. **Anais** [...]. Niterói: UFF, 2010. Disponível em: https://www.researchgate.net/profile/Martius-Rodriguez/publication/268286919_SOFTWARES_PARA_ANALISE_DE_REDES_SOCIAIS_-ARS/links/55db4a4408aec156b

[9afe676/SOFTWARES-PARA-ANALISE-DE-REDES-SOCIAIS-ARS.pdf/](#). Acesso em: 09 jun. 2023.

BALANCIERI, R. *et al.* A análise de redes de colaboração científica sob as novas tecnologias de informação e comunicação: um estudo na Plataforma Lattes. **Ciência da Informação**, Brasília, v. 34, n. 1, p. 64-77, 2005. Disponível em: <https://www.scielo.br/j/ci/a/C65dbRvBt77DQ3TQfSmDtPx/abstract/?lang=pt/>. Acesso em: 09 jun. 2023.

BARDIN, L. **Análise de conteúdo**. 6. ed. Lisboa, Portugal: EDIÇÕES 70, 2015.

BORDIN, A. S.; GONÇALVES, A. L.; TODESCO, J. L. Análise da colaboração científica departamental através de redes de coautoria. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 19, n. 2, p. 37-52, 2014. Disponível em: <https://www.scielo.br/j/pci/a/4sL4K9F59dvHnkPDKzS7YK/?lang=pt#>. Acesso em: 09 jun. 2023.

BUFREM, L. S.; GABRIEL JÚNIOR, R. F.; SORRIBAS, T. V. Redes sociais na pesquisa científica da área de ciência da informação. **DataGramZero – Revista de Informação**, v.12, n. 3, p. 1-14, 2011. Disponível em: https://www.researchgate.net/profile/Leilah-Santiago-Bufrem/publication/307638708_Redes_sociais_na_pesquisa_cientifica_da_area_de_ciencia_da_informacao/links/57ce6dc508ae83b37460ee5c/Redes-sociais-na-pesquisa-cientifica-da-area-de-ciencia-da-informacao.pdf/. Acesso em: 09 jun. 2023.

CRONIN, B. **The hand of science**: academic writing and rewards. Oxford: Scarecrow Press, 2005.

CUNHA, M. B.; AMARAL, S. A.; DANTAS, E. B. **Manual de estudo de usuários da informação**. São Paulo: Atlas, 2015.

FREEMAN, L. C. Some antecedents of social network analysis. **Connections**, v. 19, n. 1, p. 39-42, 1996. Disponível em: <https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=52dff3504f99e4dbb2f256375279048020352045/>. Acesso em: 09 jun. 2023.

GARTON, L.; HAYTHORNTHWAITE, C.; WELLMAN, B. Studying online social networks. **JMC – Journal of Computer-Mediated Communication**, v. 3, n. 1, p. 1-47, 1997. Disponível em: <http://jcmc.indiana.edu/vol3/issue1/garton.html/>. Acesso em: 09 jun. 2023.

KATZ, J. S.; MARTIN, B. R. What is Research Collaboration? **Research Policy**, v. 26, p. 1-18, 1997. Disponível em: https://www.researchgate.net/publication/4928544_What_is_Research_Collaboration/. Acesso em: 09 jun. 2023.

KOZINETS, R. V. **Netnografia**: realizando pesquisa etnográfica online. Porto Alegre: Editora Penso, 2015.

LAGO JÚNIOR, M. W. **Redes sociais informais intraorganizacionais e os processos e mudanças organizacionais**: estudo em uma empresa de tecnologia da informação. Orientador: Antonio Virgílio Bittencourt Bastos. 2005. 250 f. Dissertação (Mestrado profissional em Administração) - Núcleo de Programa de Pós-Graduação em Administração, Universidade Federal da Bahia, Salvador, 2005. Disponível em: <https://repositorio.ufba.br/handle/ri/9599/>. Acesso em: 09 jun. 2023.

LAZZARINI, S. G.; CHADDAD, F.R.; NEVES, M. F. O conceito de capital social e aplicações para desenvolvimento e estratégia sustentável. **Preços agrícolas**, Piracicaba, v. 14, n. 163, p.10-13, 2000.

LIMAS, R. F.; DANTAS, G. G. C. O uso de softwares para Análise de Redes Sociais. **Memex**, Belo Horizonte, v. 6, n. 13, 2009.

LIU, X. *et al.* Co-authorship networks in the digital library research community. **Information processing & management**, v. 41, n. 6, p. 1462-1480, 2005. Disponível em: <https://arxiv.org/pdf/cs/0502056.pdf/>. Acesso em: 09 jun. 2023.

MARTELETO, R. M. Análise de redes sociais - aplicação nos estudos de transferência da informação. **Ciência da informação**, Brasília, v. 30, n. 1, p. 71-81, 2001. Disponível em: <https://www.scielo.br/j/ci/a/6Y7Dyj4cVd5jdRkXJVxhxqN/?format=pdf&lang=pt/>. Acesso em: 09 jun. 2023.

MATHEUS, R. F.; VANZ, S. A. S.; MOURA, A. M. M. Coautoria e co-invenção: indicadores da colaboração em CT&I no Brasil. *In*: CONGRESSO IBEROAMERICANO DE INDICADORES DE CIENCIA Y TECNOLOGÍA – RICYT, 7., 2007, São Paulo. **Anais** [...]. São Paulo: 2007. p. 1-16. Disponível em: https://www.researchgate.net/publication/228584038_Co-autoria_e_co-invencao_indicadores_da_colaboracao_em_CTI_no_Brasil/. Acesso em: 09 jun. 2023.

MENA-CHALCO, J. P.; DIGIAMPIETRI, L. A.; CESAR-JÚNIOR, R. M. Caracterizando as redes de coautoria de currículos Lattes. *In*: BRAZILIAN WORKSHOP ON SOCIAL NETWORK ANALYSIS AND MINING, 1., 2012, Curitiba. **Anais** [...]. Curitiba: SBC, 2012. p. 1-12. Disponível em: <https://sol.sbc.org.br/index.php/brasnam/article/view/6864/>. Acesso em: 09 jun. 2023.

NEWMAN, M. E. J. Coauthorship networks and patterns of scientific collaboration. **PNAS**, Washington, v. 101, Suppl. 1, p. 5200-5205, 2004. Disponível em: <https://www.pnas.org/doi/epdf/10.1073/pnas.0307545100/>. Acesso em: 09 jun. 2023.

OLIVEIRA, E. F. T.; GRÁCIO, M. C. C. Rede de colaboração científica no tema “estudos métricos”: um estudo de co-autorias através dos periódicos do Scielo da área de ciência da informação. **Brazilian Journal of Information Science**, v. 2, n. 2,

p. 35-49, 2008. Disponível em:

<https://revistas.marilia.unesp.br/index.php/bjis/article/view/47/>. Acesso em: 09 jun. 2023.

ORTA, C. G. **+100 herramientas para el análisis de redes sociales #sna #ars**.

2016. Disponível em:

<https://www.k-government.com/2016/06/28/100-herramientas-analisis-redes-sna-ars/>. Acesso em: 09 jun. 2023.

PISCIOTTA, K. Redes sociais: articulação com os pares e com a sociedade. *In*:

POBLACIÓN, D. A.; WITTER, G. P.; SILVA, J. F. M. (Orgs.). **Comunicação &**

produção científica: contexto, indicadores e avaliação. São Paulo: Angellara, 2006. p.115-135.

RECUERO, R. **Introdução à análise de redes sociais online**. Salvador: EDUFBA,

2017. Disponível em: <https://repositorio.ufba.br/handle/ri/24759/>. Acesso em: 09 jun. 2023.

RECUERO, R.; BASTOS, M. T.; ZAGO, G. **Análise de Redes para Mídia Social**.

Porto Alegre: Sulina, 2015.

SILVA, A. K. A. A dinâmica das redes sociais e as redes de coautoria. **Perspectivas**

em Gestão & Conhecimento, v. 4, n. especial, p. 27-47, 2014. Disponível em:

<https://periodicos.ufpb.br/index.php/pgc/article/view/21275/>. Acesso em: 09 jun. 2023.

SILVA, T. F. *et al*. Características e especificidades da Metodologia de Análise de Redes Sociais. **Research, Society and Development**, v. 10, n. 3, p. 1-14, 2021.

Disponível em: <https://rsdjournal.org/index.php/rsd/article/view/13622/>. Acesso em: 09 jun. 2023.

SOUZA, Q. R.; QUANDT, C. O. Metodologia de análise de Redes Sociais. *In*:

DUARTE, F.; SOUZA, Q. R.; QUANDT, C. O. (Org.). **O Tempo das Redes**. São

Paulo: Perspectiva, 2008, p. 31-63. Disponível em:

https://www.academia.edu/257818/Metodologia_De_An%C3%A1lise_De_Redres_Sociais/. Acesso em: 09 jun. 2023.

VAN ECK, N. J.; WALTMAN, L. **VOSviewer manual**. Manual for VOSviewer version 1.6.19, Países Baixos: 2023. p. 54. Disponível em:

https://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.20.pdf. Acesso em: 09 jun. 2023.

VANZ, S. A. S.; STUMPF, I. R. C. Colaboração científica: revisão teórico-conceitual.

Perspectivas em Ciência da Informação, Belo Horizonte, v. 15, n. 2, p. 42- 55,

2010. Disponível em: <https://periodicos.ufmg.br/index.php/pci/article/view/23632/>.

Acesso em: 09 jun. 2023.