

QUAL FERRAMENTA BIBLIOMÉTRICA ESCOLHER? UM ESTUDO COMPARATIVO ENTRE SOFTWARES

Paulo Sergio da Conceição Moreira
Universidade Federal do Paraná
psxm54@gmail.com

André José Ribeiro Guimarães
Universidade Federal do Paraná
andrejrg@gmail.com

Denise Fukumi Tsunoda
Universidade Federal do Paraná
dtsunoda@gmail.com

Resumo

Realiza comparação entre softwares para auxiliar no processo de seleção das ferramentas adequadas para pesquisas bibliométricas. A partir de uma lista com 16 ferramentas apresentadas em publicações precedentes, seleciona quatro soluções com base em critérios de exclusão predefinidos pelos autores. A partir da exploração das funcionalidades presentes em cada uma das ferramentas, os resultados indicam que, destas quatro, o Biblioshiny apresenta o maior número de possibilidades em análises, embora apresente limitações nos relatórios visuais. O VOSviewer adota uma interface limpa e simples focada na visualização de redes bibliométricas. O Publish or Perish exibe a maior restrição em relação aos critérios comparativos, mas sua integração com o Google Acadêmico lhe confere vantagem em relação aos demais. Por fim, o CiteSpace oferece uma grande quantidade de parâmetros para formar suas visualizações de redes, mas sua interface desorganizada prejudica a experiência dos usuários. De acordo com o conteúdo analisado, constata um processo de renovação dos softwares bibliométricos. Conclui que não há uma única solução para as principais demandas da bibliometria e que a combinação entre as ferramentas tende a oferecer o melhor resultado aos pesquisadores.

Palavras-chave: Bibliometria. Comparação de ferramentas. Ferramentas bibliométricas. Redes bibliométricas.

WHICH BIBLIOMETRIC TOOL TO CHOOSE? A COMPARATIVE STUDY AMONG SOFTWARES

Abstract

This study makes a comparison between software to assist in the process of selecting the appropriate tools for bibliometric research. From a list of 16 tools presented in previous publications, four solutions are select based on exclusion criteria predefined by the authors. From the exploration of the features present in each of the tools the results indicate that, of these four, Biblioshiny presents the highest number of possibilities in analysis, although it presents limitations in the visual reports. VOSviewer adopts a clean and straightforward interface focused on the visualization of bibliometric networks. Publish or Perish has the most significant restriction on similar criteria, but its integration with Google Scholar gives it an edge over others. Subsequently, CiteSpace offers many parameters to form network views, but its disorganized interface harms the user experience. According to the content analyzed, a process of bibliometric software renovation can be realized. Finally, this paper concludes that there is not a single solution to the main demands of bibliometrics and that the combination of the tools tends to offer the best result to the researchers.

Keywords: Bibliometrics. Comparison of tools. Bibliometric tools. Bibliometric networks.



1 INTRODUÇÃO

Softwares têm sido fundamentais para a ciência. Análises, simulações, visualizações são cenários do mundo científico que são diretamente dependentes de sistemas computacionais (HOWISON *et al.*, 2015). O volume, os tipos e a granularidade de dados produzidos pelas pesquisas científicas fazem com que o software assuma papel chave no fluxo de trabalho dos pesquisadores das mais diversas áreas do conhecimento (BORGMAN; WALLIS; MAYERNIK, 2012).

Uma das áreas que mais se beneficia da adoção dos softwares acadêmicos é a bibliometria. As análises bibliométricas têm por objetivo a aplicação de métodos matemáticos e estatísticos às publicações acadêmicas (DIODATO; GELLATLY, 2013; THOMPSON; WALKER, 2015) de forma a permitir a avaliação de padrões relacionados à publicação e à utilização de documentos (DIODATO; GELLATLY, 2013). Além disso, as análises de cunho bibliométrico possibilitam a mensuração e a avaliação de resultados de pesquisas apresentados na forma escrita, nas mais diversas formas de mídia (BALL, 2017). Esse vasto campo de trabalho apresenta desafios que demandam grande esforço no projeto e na implementação de softwares bibliométricos (HOWISON; BULLARD, 2016).

Ainda assim, mesmo com o aumento da disponibilidade e da variedade de softwares voltados às pesquisas acadêmicas, incluindo a bibliometria, sua relevância em publicações científicas foi por muito tempo diminuída ou mesmo ignorada (PAN *et al.*, 2018). Howison e Bullard (2016) afirmam que, no entanto, a visibilidade do software na ciência está em questão, especialmente pela necessidade em entender e confirmar a qualidade pesquisas realizadas ou em andamento. Esta preocupação com a qualidade das pesquisas, em relação ao software, extrapola questões técnicas e considera elementos que visam agilizar o trabalho dos pesquisadores como compartilhamento de dados e recursos.

Neste contexto, cuja utilização de softwares mostra-se essencial à bibliometria - mesmo ainda tendo sido pouco explorada pelas pesquisas científicas - o objetivo do presente artigo é analisar o potencial de diferentes ferramentas bibliométricas. Para isso, realiza-se um processo de comparação entre quatro softwares, a saber: a) Biblioshiny; b) CiteSpace; c) Publish or Perish; e d) VOSViewer. Estas ferramentas foram selecionadas por atenderem aos seguintes que atendem aos seguintes critérios predefinidos pelos autores: a) versão recente; b) interface gráfica; e c) que sejam de livre utilização. Dessa forma, este comparativo apresenta as vantagens, desvantagens e principais diferenças entre as ferramentas estudadas, apresentando as situações mais adequadas para seu emprego.

2 REFERENCIAL TEÓRICO

A bibliometria viabiliza a identificação dos progressos realizados em determinada área do conhecimento; o reconhecimento de fontes fidedignas de publicação científica; o estabelecimento de uma base consistente para a avaliação de trabalhos científicos; o desenvolvimento de índices bibliométricos para avaliação da produção acadêmica; entre outros indicadores (DAIM *et al.*, 2006; THELWALL, 2008; MARTÍNEZ *et al.*, 2015).

Ademais, cabe ressaltar que a bibliometria favorece o acesso e a análise de pesquisas realizadas pelos mais distintos atores, como países, universidades, centros e grupos de pesquisa, revistas científicas e pesquisadores (MARTÍNEZ *et al.*, 2015), possibilitando aos pesquisadores o conhecimento, por exemplo, da estrutura cognitiva dos campos de pesquisa, o surgimento de novas frentes de pesquisa ou padrões nacionais e internacionais de coautoria (THELWALL, 2008).

Para Noyons, Moed e Van Raan (1999) e Moed, Glänzel e Schmoch (2004), os dois principais métodos bibliométricos para exploração da produção científica sobre um campo de pesquisa são a análise de desempenho e o mapeamento científico (*performance analysis* e *science mapping*), que podem ser realizadas individualmente ou em conjunto.

A análise de desempenho tem por objetivo avaliar atores científicos ou grupos deles, baseando-se em dados bibliográficos que indicam a atividade de países, universidades, faculdades, departamentos, e o impacto de sua produção. O desempenho é mensurado por meio do número de publicações em um período predeterminado, enquanto que o impacto é aferido pelo número de vezes em que as publicações são citadas por outros (NOYONS; MOED; VAN RAAN, 1999).

Os mesmos autores definem que o mapeamento científico, por sua vez, é orientado ao monitoramento de áreas de pesquisa. Desta forma, é possível identificar tendências em campos científicos, sua evolução, seus *clusters* e qual a interação entre estes. Com o mapeamento científico, também é possível realizar o monitoramento da evolução temporal e reconhecer os principais autores envolvidos (NOYONS; MOED; VAN RAAN, 1999).

Para a realização do mapeamento científico são utilizados mapas dinâmicos, em que se pode observar o crescimento ou a diminuição da produção, bem como a interação entre os subdomínios (NOYONS; MOED; VAN RAAN, 1999). A combinação do mapeamento científico com métodos bibliométricos permite, por exemplo, em comparação a uma revisão estruturada da literatura, análises capazes de lidar com uma ampla gama de centenas, até milhares, de estudos (ZUPIC; ČATER, 2015).

Os dois métodos bibliométricos têm em comum a necessidade da realização de uma etapa pré-processamento dos dados para a posterior compreensão dos resultados. Cobo *et al.* (2011) defendem que a etapa de pré-processamento talvez seja a mais importante em um estudo bibliométrico, visto que a confiabilidade de seus resultados irá depender da qualidade dos dados. Dentre as fases de pré-processamento, podem-se citar a detecção de duplicações ou elementos com erros de escrita (COBO *et al.*, 2011).

Tendo em vista este argumento, bem como a necessidade de se “padronizar” uma análise bibliométrica, Zupic e Čater (2015) propuseram um fluxo de trabalho para pesquisas baseadas em métodos bibliométricos. O fluxo em questão é composto por cinco etapas, a saber: a) definição da questão de pesquisa e escolha dos métodos bibliométricos adequados para as respostas; b) seleção das bases de dados; c) utilização de um software bibliométrico para as análises; d) decisão sobre qual método de visualização utilizar para representar os resultados obtidos no terceiro passo; e e) interpretação e relato dos resultados obtidos.

No que tange principalmente aos passos c e d do fluxo proposto Zupic e Čater (2015), diversas ferramentas têm sido desenvolvidas (VAN ECK; WALTMAN, 2009; PERSSON; DANELL; SCHNEIDER, 2009; COBO *et al.*, 2011, 2012; MCLEVEY; MCILROY-YOUNG, 2017; ARIA; CUCCURULLO, 2017), cada qual com características e objetivos distintos, que devem ser conhecidos para que seja possível a seleção da ferramenta mais apropriada para cada finalidade.

143

3 PROCEDIMENTOS METODOLÓGICOS

Trata-se de uma pesquisa descritiva de abordagem qualitativa, cujo primeiro procedimento adotado correspondeu à identificação de um conjunto de ferramentas bibliométricas para efeitos de comparação. Nesta etapa, tomou-se como base os artigos elaborados por Cobo *et al.* (2011, 2012) e McLevey e McIlroy-Young (2017), por serem trabalhos que apresentavam comparações entre ferramentas bibliométricas e a lista de softwares bibliométricos e de colaboração científica presente no *Portal da Escrita Científica*, da Universidade de São Paulo (USP), campus São Carlos.

Após esta etapa, 16 ferramentas foram pré-selecionadas, conforme elenca-se no Quadro 1.

Quadro 1- Lista das ferramentas pré-selecionadas para análise

Ferramenta	Referência
Bibexcel	Persson, Danell e Schneider (2009)
Bibliometrix/Biblioshiny	Aria e Cuccurullo (2017)
CiteSpace	Chen (2006)
CoPalRed	Bailón-Moreno, Jurado-Alameda e Ruiz-Baños (2006)
IN-SPIRE™	Pacific Northwest National Laboratory (2019)
InCites	Clarivate Analytics (2019)
Leydesdorff's Software	-
Metaknowledge	Mclevey e Mcilroy-Young (2017)
Network Workbench Tool	NWB Team (2006)
Publish or Perish	Harzing (2007)
Science of Science (Sci ²) Tool	Sci ² Team (2009)
SCImago	SCImago (2019)
SciMAT	Cobo <i>et al.</i> (2012)
SciVal	Elsevier (2019)
VantagePoint	Porter e Cunningham (2005)
VOSViewer	Van Eck e Waltman (2009)

Fonte: Elaborado pelos autores (2019).

Das ferramentas elencadas no Quadro 1, duas foram excluídas diretamente: a) Leydesdorff's Software, por ser tratar de um conjunto de ferramentas desenvolvidas pelo pesquisador Loet Leydesdorff, em que cada ferramenta apresenta objetivos diferentes; e b) CoPalRed (BAILÓN-MORENO; JURADO-ALAMEDA; RUIZ-BAÑOS, 2006), uma vez que o site da ferramenta estava indisponível, não sendo possível ter contato com a ferramenta.

Tendo em vista a grande quantidade de ferramentas identificadas, realizou-se uma “triagem” para reduzir o número de ferramentas. Desta forma, uma lista de critérios foi estabelecida empiricamente pelos autores para esta etapa. Os critérios de exclusão adotados são elencados no Quadro 2.

Quadro 2 - Critérios de exclusão

Critérios	Descrição	Ferramentas excluídas
Licença	Foram mantidas apenas ferramentas gratuitas	VantagePoint; SciVal; InCites; IN-SPIRE™
Versão atualizada	Optou-se por analisar apenas as ferramentas cuja última atualização foi no período de 2017 a 2019	SciMAT; Network Workbench Tool
Qualidade da documentação	Ferramentas cuja documentação foi considerada “baixa” foram eliminadas	Bibexcel; Network Workbench Tool; Science of Science (Sci ²) Tool; InCites
Interface gráfica	Optou-se por manter apenas as ferramentas com interface gráfica	Metaknowledge; Bibliometrix*

Fonte: Elaborado pelos autores (2019).

Quanto à qualidade da documentação, estabeleceu-se uma nova lista de critérios para realizar a comparação: a) documentação oficial, fornecida pelo desenvolvedor; b) documentação recente, atualizada entre 2017 e 2019; c) explicação do significado de cada métrica; d) detalhe dos parâmetros das análises; e) existência de exemplos de como realizar as análises; e f) existência de outras fontes de documentação (notebooks; slides; vídeos, tutoriais).

Com base nestes critérios, utilizou-se os termos “baixa”, “média” e “alta” para a qualidade da documentação, sendo que considerou-se uma documentação com qualidade “alta” aquela que atendessem a todos os critérios estabelecidos. Caso um dos critérios não fosse atendido, a documentação foi considerada como “média” e quando dois ou mais critérios não fossem atendidos, a qualidade da documentação foi considerada “baixa”.

Finalmente, com relação à interface gráfica, todas as ferramentas elencadas no Quadro 3 apresentam este tipo de recurso, com exceção às ferramentas Metaknowledge e Bibliometrix. O pacote Bibliometrix - que é por padrão uma interface de linhas de comando - oferece uma interface gráfica denominada Biblioshiny possibilitando sua utilização sem conhecimento em programação. Sendo assim, as análises foram destinadas a esta interface e não à ferramenta executada por scripts.

Uma síntese das considerações expostas acima é apresentada no Quadro 3.

Quadro 3 - Ferramentas analisadas

Ferramenta	Gratuita	Última Versão	Qualidade da documentação	Interface gráfica?
Bibliometrix/Biblioshiny	Sim	14/03/2019	Alta	Parcial
VOSViewer	Sim	03/04/2019	Alta	Sim
Publish or Perish	Sim	17/04/2019	Alta	Sim
CiteSpace	Sim	31/08/2018	Média	Sim
Metaknowledge	Sim	21/01/2019	Média	Não
Bibexcel	Sim	2017	Baixa	Sim
Network Workbench Tool	Sim	15/09/2009	Baixa	Sim
Science of Science (Sci ²) Tool	Sim	31/01/2018	Baixa	Sim
InCites	Não	N/A	Baixa	Sim
SciMAT	Sim	12/07/2016	Média	Sim
IN-SPIRE™	Não	01/02/2019	Média	Sim
VantagePoint	Não	05/02/2019	Média	Sim
SciVal	Não	26/03/2019	Média	Sim
SCImago	Sim	N/A	Média	Sim

Fonte: Elaborado pelos autores (2019).

Conforme os critérios estabelecidos, quatro ferramentas foram selecionadas para a fase de avaliação: a) Biblioshiny; b) CiteSpace; c) Publish or Perish e d) VOSViewer. Para a análise destas ferramentas, testou-se cada ferramenta com o intuito de verificar as suas funcionalidades e limitações. Além disso, recorreu-se à leitura das respectivas documentações e materiais de suporte de cada software, como elemento complementar.

146

A seguir, apresenta-se uma breve descrição de cada uma destas ferramentas, bem como os respectivos resultados obtidos.

4 ANÁLISE E DISCUSSÃO DOS DADOS

4.1 Biblioshiny

Trata-se de um pacote desenvolvido para a linguagem R e que fornece um conjunto de ferramentas para pesquisas relacionadas à bibliometria e à cientometria (ARIA; CUCCURULLO, 2017). Consiste em uma ferramenta de código aberto que abrange uma gama de análises quantitativas e encontra-se disponível para os sistemas operacionais Linux; Mac OS X; Windows e pode ser utilizada via comandos (scripts) ou por meio de uma interface gráfica. A ferramenta encontra-se disponível em: <http://www.bibliometrix.org/biblioshiny.html>.

Em termos de bases de dados, a interface Biblioshiny permite a análise de registros oriundos das bases Web of Science (WOS) e Scopus. Para os registros da WOS, é possível analisar aqueles nos formatos “.txt”, “.bib”, “.xlsx” e “.RData files”. Também é possível

analisar múltiplos arquivos “.txt” ou “.bib”, desde que estes estejam compactados em arquivos “.zip”.

No que tange às análises “básicas”, esta ferramenta permite avaliar três grandes grupos: a) autores, incluindo análises relacionadas à autoria, à afiliação e aos países; b) fontes de publicações, permitindo a avaliação do impacto das fontes e verificação da produtividade; e c) documentos, englobando referências e palavras mais citadas.

O conjunto de análises disponíveis na ferramenta é apresentado no Quadro .

Quadro 4 - Conjunto de análises disponíveis no Biblioshiny

Autores	Autores: Autores mais relevantes (Produção); Produção ao longo dos anos; Lei de Lotka; Impacto (<i>H-Index</i> ; <i>G-Index</i> ; <i>M-Index</i> ; Total de citações); Afiliação: Afiliações mais relevantes; País: País do autor; Países mais citados; Produção científica por país.
Fontes	Fontes mais relevantes (produção); Fontes mais citadas; Lei de Bradford; Impacto (<i>H-Index</i> ; <i>G-Index</i> ; <i>M-Index</i> ; Total de citações); <i>Source Dynamics</i>
Documentos	Documentos: Documentos mais citados (global) Documentos mais citados (local); Referências: Referências mais citadas; <i>Reference Spectroscopy</i> ; Palavras: Palavras mais citadas (Palavras-chave (autor); <i>Keyword Plus</i> ; Título; Resumo); <i>Word Dynamics</i> ; <i>WordCloud</i> ; <i>TreeMap</i>

Fonte: Elaborado pelos autores (2019).

Destaca-se a capacidade da ferramenta em fornecer diretamente os resultados para as leis de Lotka (LOTKA, 1926) e Bradford (BRADFORD, 1985), além dos índices H (HIRSCH, 2005) e G (EGGHE, 2006). Todos os resultados são fornecidos em formato de tabelas e podem ser exportados para “CSV” e “PDF”, além de ser possível exportar diretamente para planilhas Excel.

Por fim, em relação aos elementos de visualização, o Biblioshiny apresenta, basicamente: a) gráficos de barras, para análises que envolvem contagens e b) gráfico de linhas, para as análises que correspondem à evolução da produção ao longo dos anos. A ferramenta também dispõe de elementos gráficos específicos para análises de palavras, como nuvem de palavras (*word cloud*) e *treemap*.

Além dos recursos dispostos no Quadro , é possível realizar análises para três tipos de estruturas: a) estrutura conceitual, relacionada às palavras e que envolve análises de coocorrência e fatorial, mapa temático e evolução do tema; b) estrutura intelectual, destinada aos trabalhos (*papers*), aos autores e às fontes, englobando rede de cocitação e historiografia; e c) estrutura social, que objetiva a demonstrar a interação de autores, instituições e países.

A síntese das análises disponíveis em termos de estruturas é apresentada no Quadro .

Quadro 5 - Análises de estruturas

Estrutura	Análise	Campos	Parâmetros
Conceitual	Rede de coocorrência	Keywords Plus; Palavras-chave (autor); Título; Resumo	Normalização: <i>Association; Jaccard; Salton; Inclusion; Equivalence.</i> Agrupamento: <i>Edge Betweenness; InfoMap; Leading Eigenvalues; Louvain; Spinglass; Walktrap.</i>
	Mapa temático		N/A
	Evolução do tema		N/A
	Análise fatorial		Análise de correspondência; Análise de correspondência múltipla; Escalonamento Multidimensional.
Intelectual	Rede de cocitação	Trabalhos (<i>papers</i>); Autores; Fontes	Agrupamento: <i>Edge Betweenness; InfoMap; Leading Eigenvalues; Louvain; Spinglass; Walktrap.</i>
	Historiografia	N/A	Algoritmo de busca: <i>Fast search; Full search</i>
Social	Rede de colaboração	Autores; Instituições; Países	Agrupamento: <i>Edge Betweenness; InfoMap; Leading Eigenvalues; Louvain; Spinglass; Walktrap.</i>
	Mapa de colaboração entre países	N/A	N/A

Fonte: Elaborado pelos autores (2019).

Para cada análise disposta no Quadro , apresenta-se os respectivos parâmetros associados e que podem influenciar o resultado desejado pelo pesquisador.

4.2 CiteSpace

Segundo o manual do CiteSpace, traduzido pelo Grupo de Pesquisa de Recuperação da Informação e Tecnologias Avançadas (RITA) da Universidade Federal de Santa Catarina, o software foi “projetado para responder perguntas sobre o domínio do conhecimento” (CHEN, 2016, p. 6). O CiteSpace fundamenta-se na análise e na visualização de redes que possibilitam a exploração da paisagem intelectual de uma área de investigação ou disciplina científica. Por meio de algoritmos computacionais que resultam em uma interface interativa, o CiteSpace configura-se em uma ferramenta de rastreamento do desenvolvimento da ciência revelado pelas publicações acadêmicas. A ferramenta está disponível em: <http://cluster.cis.drexel.edu/~cchen/citespace>.

O CiteSpace foi desenvolvido na linguagem Java, atendendo aos principais sistemas operacionais vigentes (Windows, Mac OS X e Linux). Lançado em 2003, a versão mais recente do programa (5.3.R4) é datada de 31 de agosto de 2018 e suporta arquivos das bases de dados WOS, Scopus, CSCD, CSSCI e Pubmed. Para todas estas opções, os arquivos

importados devem estar no padrão de texto, sem formatação e, obrigatoriamente, devem possuir nomes iniciando com a expressão “download”.

O CiteSpace trabalha com o conceito de projeto, onde é necessário definir um diretório para que os arquivos gerados sejam salvos e outro diretório correspondente à fonte dos dados importados.

Para a formação do mapa, o CiteSpace permite a seleção de um ou mais tipos de nós que compõem a rede. Por padrão, a rede é formada pelas referências citadas, gerando uma análise de cocitação, e ao se selecionar outros valores, há a formação de redes de tipos diferentes (colaboração, cocitação e coocorrência). Os parâmetros possíveis de serem combinados são “*Author*”, “*Institution*”, “*Country*”, “*Term*”, “*Keyword*”, “*Source*”, “*Category*”, “*Reference*”, “*Cited Author*” e “*Cited Journal*”. Todavia, salienta-se que nem todas combinações são permitidas, sendo que o próprio software se encarrega de limitar as configurações possíveis.

Outro aspecto importante é o critério de seleção dos nós da rede. O padrão é a escolha das 50 ocorrências mais citadas (valor que pode ser alterado pelo usuário), mas pode-se definir como critério uma porcentagem dos registros mais citados, o índice G (EGGHE, 2006), a interpolação limiar, um intervalo específico de citação e os documentos mais citados, segundo da WOS, dos últimos 180 dias ou dos últimos cinco anos.

149

4.3 Publish or Perish

O Publish or Perish (HARZING, 2007) é um software gratuito capaz de recuperar e analisar citações acadêmicas de diferentes fontes, incluindo o Google Acadêmico e a Microsoft Academic Research. Há versões disponíveis para GNU Linux; OS X e Windows. O Publish or Perish possui integração com as bases Crossref, Google Scholar, Microsoft Academic, WOS e Scopus. Além disso, apresenta compatibilidade com arquivos CSV (Publish or Perish, Google Scholar/Citations e Scopus), RefMan/RIS, EndNote e todos os formatos disponibilizados pela WOS, com exceção aos arquivos para MAC.

Entre as análises básicas oferecidas pelo Publish or Perish, estão o número de trabalhos, total de citações, média de citações por trabalho, número de citações por autor (com possibilidade de filtragem por ano), número de trabalhos por autor e média de autores por trabalho. O software também fornece indicadores de impacto como índices H (HIRSCH, 2005) e G (EGGHE, 2006), além de algumas variações como índice H contemporâneo, índice H individual, média anual do aumento do índice H e taxa de citação ponderada por idade.

O diferencial desta ferramenta consiste na capacidade de explorar uma grande variedade de arquivos oriundos de diferentes bases científicas e as variações de análises para se medir o impacto dos autores. A ferramenta está disponível em: <https://harzing.com/resources/publish-or-perish>.

4.4 VOSviewer

O VOSviewer é um software orientado à criação, visualização e exploração mapas baseados em dados de redes (VAN ECK; WALTMAN, 2018). Projetado inicialmente para analisar dados bibliométricos, o VOSviewer pode ser utilizado em qualquer conjunto de dados de redes e adota o método conhecido como VOS (*Visualization of Similarities*) para definir os nós e ligações de sua rede. Encontra-se disponível para download em: <http://www.vosviewer.com/>.

O método VOS tem como objetivo principal construir visualizações em duas dimensões nas quais os objetos com alta similaridade são localizados mais proximamente que objetos com baixa ou nenhuma similaridade (VAN ECK; WALTMAN, 2007). Em redes bibliométricas, se dois pesquisadores estão proximamente localizados na visualização, maior a ligação entre eles, ou seja, maior a tendências de eles serem citados nas mesmas publicações (VAN ECK; WALTMAN, 2014).

Adicionalmente, o VOSviewer permite a visualização de *clusters*, grupos de itens com características em comum no mapa (VAN ECK; WALTMAN, 2018).

Desenvolvido em Java, o VOSviewer possui versões específicas para Mac OS X e Windows, mas também é suportado por outros sistemas operacionais. Por meio de uma interface exclusivamente gráfica para a criação de redes, fornece suporte a arquivos com dados bibliográficos gerados pelas bases de dados WOS, Scopus, Dimensions e PubMed. Além disso, fornece suporte a arquivos procedentes de programas de gerenciamento de referências (RIS, EndNote e RefWorks) e obtém dados de API (*Application Programming Interface*) externa (Crossref, Europe PMC, Semantic Scholar, OCC, COCI e Wikidata). Para esta segunda opção, a busca pode ser realizada por meio de um formulário ou por um arquivo de texto que contenha os números de identificação (*Digital Object Identifier*) dos documentos a serem analisados. Ademais, o software reconhece arquivos de redes gerados externamente (GML e Pajek) e arquivos de texto plano para criação de redes entre termos de um determinado corpus.

No Quadro , são apresentados os tipos e unidades de análise disponíveis no software.

Quadro 6 - Conjunto de análises disponíveis no VOSviewer

Tipo	Unidade de análise
Coautoria	Autores; Organizações; Países
Coocorrência	Todas as palavras-chave; Palavras-chave do autor; <i>Keywords Plus (WOS)</i>
Citação	Documentos; Fontes; Autores; Organizações; Países
Pares bibliográficos	Documentos; Fontes; Autores; Organizações; Países
Cocitação	Referências citadas; Fontes citadas; Autores citados

Fonte: Elaborado pelos autores (2019).

Na criação de mapas bibliométricos, o VOSviewer adota um fluxo de trabalho onde as escolhas feitas pelo usuário determinam as opções exibidas na etapa seguinte. Este processo faz com que as telas apresentadas sejam mais simples, diminuindo a possibilidade de erros. Todavia, ressalta-se que nem todas as análises apresentadas estão disponíveis a todos os modos de entrada de dados. A presença das referências utilizadas pelos documentos analisados é determinante para a análise de cocitação, por exemplo.

Com o mapa gerado, o VOSviewer apresenta, além das opções para ajustes na visualização, uma área com os itens que compõem os *clusters* e um painel com os parâmetros utilizados na formação da rede. Destacam-se os métodos de normalização dos dados (sem normalização, associação, fracionalização e modularidade LinLog) (VAN ECK; WALTMAN, 2018), os parâmetros de atração e repulsão dos nós da rede e, as definições acerca da formação dos *clusters*. Por fim, o resultado pode ser exportado para diversos formatos de imagem ou ser salvo para futura utilização.

4.5 Comparação entre as ferramentas

Considerando as etapas do fluxo de trabalho de uma análise bibliométrica proposto por Zupic e Čater (2015), julga-se que as ferramentas analisadas podem ser empregadas majoritariamente nas etapas três - utilização de software bibliométrico para análises -, e quatro - visualização para representar os resultados.

Além disso, a segunda etapa, que corresponde à seleção das bases de dados, é contemplada por todas ferramentas estudadas. Embora, em muitos momentos, não seja possível distinguir com clareza os procedimentos das etapas três e quatro, optou-se por distinguir os critérios de comparação entre as ferramentas entre estes dois grupos. No

Quadro , são dispostas as ferramentas e os componentes de avaliação, onde o “X” indica que o item referido foi atendido.

Quadro 7 - Comparativo entre as ferramentas

			Biblioshiny	VOSviewer	Publish or Perish	CiteSpace
Análises básicas	Autores	Autores mais relevantes	X	X	X	X
		Produção ao longo dos anos	X			X
		Lei de Lotka	X			
		Índices de impacto (H-Index, G-Index, M-Index)	X		X	
		Total de citações	X	X	X	X
		Afiliações mais relevantes	X	X		X
		País do autor	X			
		Países mais citados	X	X		
		Produção científica por país	X	X		X
	Fontes	Fontes mais relevantes	X	X		X
		Fontes mais citadas	X	X		
		Lei de Bradford	X			
		Índices de impacto (H-Index, G-Index, M-Index)	X			
		Total de citações	X	X		
		Source Dynamics	X			
	Documentos	Documentos mais citados	X	X	X	
		Referências mais citadas	X	X		X
		Palavras mais citadas	X	X		
		Keyword Plus	X	X		X
		Word Dynamics	X			
		WordCloud	X	X		X
TreeMap		X				
Visualização	Coautoria	Autores	X	X		
		Organizações	X	X		X
		Países	X	X		X
	Coocorrência	Todas as palavras-chave	X	X		X
		Palavras-chave do autor	X	X		X
		Keyword Plus (WoS)	X	X		X
	Citação	Documentos	X	X		X
		Fontes	X	X		X
		Autores	X	X		X
		Organizações	X	X		
		Países	X	X		
	Pares bibliográficos	Documentos	X	X		
		Fontes	X	X		
		Autores	X	X		
		Organizações	X	X		
		Países	X	X		
	Cocitação	Referências citadas	X	X		X
Fontes citadas		X	X		X	
Autores citados		X	X		X	

Fonte: Elaborado pelos autores (2019) com base em Cobo *et al.* (2011, 2012).

Como se observa no Quadro 7, o Biblioshiny destaca-se como a ferramenta com o maior conjunto de análise, atendendo a todos os critérios de comparação estipulados. Ela se diferencia das demais por ser a única a não ter sido desenvolvida na linguagem Java e por utilizar uma interface web fundamentada na linguagem R. Esse fator pode ser visto como vantagem, uma vez que permite o maior engajamento da comunidade e utilização dos pacotes

R; entretanto, ao mesmo tempo, apresenta-se como uma possível limitação, especialmente nas análises gráficas. O Biblioshiny apresenta praticamente todas as análises presentes no VOSviewer, porém conta com severas restrições de personalização no visual das redes formadas, sendo indicada a exportação dos dados para outras ferramentas para um resultado mais apurado.

Em contrapartida, o VOSviewer não fornece análises específicas para as leis de Lotka (LOTKA, 1926) e Bradford (BRADFORD, 1985), ou para os índices H (HIRSCH, 2005) e G (EGGHE, 2006). Uma vez que o foco do software é o método VOS, seu ponto forte se concentra na visualização de redes. Ainda assim, antes da visualização do mapa, é possível verificar os dados processados, em formato de tabela que pode ser exportado para outros programas, possibilitando novos tipos de análise. De qualquer maneira, o VOSviewer se caracteriza como uma ferramenta típica da etapa de visualização dos resultados.

O software Publish or Perish foi o que menos critérios atendeu dentro as ferramentas avaliadas. Seu foco consiste na análise de citações e variações destas. A possibilidade de realizar análises diretamente a partir do Google Acadêmico é seu diferencial positivo, mostrando-se útil à etapa de seleção de bases de dados. Todavia, o fato de não oferecer recursos de visualização impede sua utilização na quarta etapa do processo.

Embora a fundamentação do CiteSpace se alinhe à adotada pelo VOSviewer, as abordagens das duas ferramentas se distinguem sensivelmente. Entre as diferenças que podem ser consideradas positivas ao CiteSpace, destaca-se a interatividade com os elementos do mapa após sua formação inicial, sendo possível arrastar os elementos para uma melhor disposição. É possível também a evidenciação dos *clusters* e sua identificação, o que facilita a visualização das informações. Outros recursos positivos são a possibilidade de recorte temporal dentro da própria interface e a visualização de uma linha do tempo dos documentos analisados. Além disso, o CiteSpace oferece uma gama de parâmetros muito superior ao VOSviewer.

Por outro lado, essa grande quantidade de opções é apresentada simultaneamente ao usuário por meio de diversos painéis e abas, o que forma uma interface pouco amigável. Os mapas gerados por padrão também causam dificuldade na leitura dos rótulos dos nós da rede, atrapalhando a identificação dos elementos. A citada interatividade não é fluida - tampouco intuitiva - onde a ampliação e deslocamento do mapa na tela só são realizados diante de muito esforço. Outro aspecto que não se mostra eficiente é a adoção de ícones na personalização do mapa. Os elementos gráficos são pouco intuitivos e só há reforço textual quando o mouse está sobre o botão. Por fim, mais um fato negativo ocorreu durante o estudo. Durante o processo

de download do CiteSpace, o Windows Defender, instalado por padrão no Windows 10 Pro, estava com a opção “Proteção em tempo real” ativada e considerou o arquivo baixado como potencial ameaça à segurança do sistema, bloqueando a instalação. Foi necessário desativar esta opção para prosseguir o processo.

5 CONSIDERAÇÕES FINAIS

Definir qual ferramenta adotar em uma análise bibliométrica não é uma tarefa simples. Há muitas opções com diferentes abordagens e tecnologias que impõem ao pesquisador a pergunta: qual ferramenta bibliométrica escolher para minha pesquisa? Este artigo objetivou analisar o potencial de diferentes ferramentas bibliométricas. Para isso, identificou-se 16 ferramentas bibliométricas para auxiliar na busca por responder esta dúvida. Dentre essas, quatro são selecionadas para uma aprofundada comparação de recursos, ressaltando as vantagens e deficiências de cada uma.

A Biblioshiny, interface gráfica do pacote R Bibliometrix, mostrou-se a mais completa das ferramentas analisadas. Possui grande abrangência no número de análises, mas apresentou deficiência em relação à parte de visualização. Ressalta-se, porém, que por ser uma ferramenta lançada em 2017 apresenta grande potencial de evolução, valendo-se das possibilidades oriundas da linguagem R. O VOSviewer, software focado em visualização de redes bibliométrica, apresenta uma interface gráfica limpa e simples, atendendo aos principais requisitos da área.

Além disso, vem agregando novas funcionalidades constantemente, como a busca interna por meio de APIs externas, dispensando a busca por dados fora do software. O Publish or Perish, terceira ferramenta analisada, apresenta um foco muito definido onde seu principal diferencial é a integração com o Google Acadêmico. No entanto, a ausência de relatórios visuais fez com que fosse a ferramenta com menos recursos dentre as estudadas. Por fim, o CiteSpace, ferramenta de visualização de redes lançada em 2003, destaca-se pela grande quantidade de parâmetros e personalização dos mapas gerados. Em contrapartida, a interface prejudica a experiência do usuário, causando-lhe desvantagem em relação às demais opções.

Diante da avaliação e comparação realizada, conclui-se que cada ferramenta pode se mostrar adequada a uma situação específica. Uma recomendação para pesquisas futuras seria um estudo aplicando as ferramentas analisadas a uma mesma base de conhecimento. Assim, diante de conteúdo comum, outros critérios poderiam ser estipulados. De qualquer maneira, é

possível perceber que os softwares bibliométricos têm passado por um processo de renovação. Adequadas às tecnologias vigentes, novas soluções estão surgindo, criando mais possibilidades e mecanismos de análises aos pesquisadores.

REFERÊNCIAS

- ARIA, M.; CUCCURULLO, C. Bibliometrix: An R-tool for comprehensive science mapping analysis. **Journal of Informetrics**, v. 11, n. 4, p. 959–975, 2017. Disponível em: <https://doi.org/10.1016/j.joi.2017.08.007>.
- BAILÓN-MORENO, R.; JURADO-ALAMEDA, E.; RUIZ-BAÑOS, R. The scientific network of surfactants: Structural analysis. **Journal of the American Society for Information Science and Technology**, v. 57, n. 7, p. 949–960, 2006.
- BALL, R. **An Introduction to Bibliometrics: New Development and Trends**. [s.l.] Chandos Publishing, 2017.
- BORGMAN, C. L.; WALLIS, J. C.; MAYERNIK, M. S. Who's got the data? Interdependencies in science and technology collaborations. **Computer Supported Cooperative Work**, v. 21, n. 6, p. 485–523, 2012.
- BRADFORD, S. Specific subjects. **Journal of Information Science**, v. 10, p. 173–180, 1985.
- CHEN, C. CiteSpace II: Detecting and visualizing emerging trends and transient patterns in scientific literature. **Journal of the American Society for information Science and Technology**, v. 57, n. 3, p. 359–377, 2006.
- CHEN, C. **The CiteSpace Manual**. 2016. Disponível em: <http://rita.ufsc.br:8080/rita/tutoriais/manual-do-citespace-em-portugues-brasileiro/view>. Acesso em: 4 maio. 2019.
- CLARIVATE ANALYTICS. **InCites**. Disponível em: <https://incites.thomsonreuters.com/>. Acesso em: 4 maio. 2019.
- COBO, M. J.; LÓPEZ-HERRERA, A. G.; HERRERA-VIEDMA, E.; HERRERA, F. Science mapping software tools: Review, analysis, and cooperative study among tools. **Journal of the American Society for Information Science and Technology**, v. 62, n. 7, p. 1382–1402, 2011.
- COBO, M. J.; LÓPEZ-HERRERA, A. G.; HERRERA-VIEDMA, E.; HERRERA, F. SciMAT: A new science mapping analysis software tool. **Journal of the American Society for Information Science and Technology**, v. 63, n. 8, p. 1609–1630, 2012.
- DAIM, T. U. et al. Forecasting emerging technologies: Use of bibliometrics and patent analysis. **Technological Forecasting and Social Change**, v. 73, n. 8, p. 981-1012, 2006.
- DIODATO, V. P.; GELLATLY, P. **Dictionary of bibliometrics**. Routledge, 2013.
- EGGHE, L. Theory and practise of the g-index. **Scientometrics**, v. 69, n. 1, p. 131–152, 2006.
- ELSEVIER. **SciVal**. Disponível em: <https://www.elsevier.com/solutions/scival>. Acesso em: 4 maio. 2019.
- P2P & INOVAÇÃO, Rio de Janeiro, v. 6 n. 2, Ed. Especial, p.140-158, 2020.

HARZING, A. W. **Publish or Perish**. Disponível em: <https://harzing.com/resources/publish-or-perish>. Acesso em: 4 maio. 2019.

HIRSCH, J. E. An index to quantify an individual's scientific research output. **Proceedings of the National academy of Sciences**, v. 102, n. 46, p. 16569–16572, 2005.

HOWISON, J.; BULLARD, J. Software in the scientific literature: Problems with seeing, finding, and using software mentioned in the biology literature. **Journal of the Association for Information Science and Technology**, v. 67, n. 9, p. 2137–2155, set. 2016.

HOWISON, J.; DEELMAN, E.; MCLENNAN, M. J.; DA SILVA, R. F.; HERBSLEB, J. D. Understanding the scientific software ecosystem and its impact: Current and future measures. **Research Evaluation**, v. 24, n. 4, p. 454–470, 2015.

LOTKA, A. J. The frequency distribution of scientific productivity. **Journal of the Washington academy of sciences**, v. 16, n. 12, p. 317–323, 1926.

MARTÍNEZ, M. A.; COBO, M. J.; HERRERA, M.; HERRERA-VIEDMA, E. Analyzing the scientific evolution of social work using science mapping. **Research on Social Work Practice**, v. 25, n. 2, p. 257–277, 2015.

MCLEVEY, J.; MCILROY-YOUNG, R. Introducing metaknowledge: Software for computational research in information science, network analysis, and science of science. **Journal of Informetrics**, v. 11, n. 1, p. 176–197, 2017.

MOED, H. F.; GLÄNZEL, W.; SCHMOCH, U. Handbook of quantitative science and technology research. In: **The Use of Publication and Patent Statistics in Studies of S&T Systems**. Springer, 2004.

NOYONS, E.; MOED, H.; VAN RAAN, A. Integrating research performance analysis and science mapping. **Scientometrics**, v. 46, n. 3, p. 591–604, 1999.

NWB TEAM. **Network Workbench Tool**. Disponível em: <http://nwb.slis.indiana.edu>. Acesso em: 4 maio. 2019.

PACIFIC NORTHWEST NATIONAL LABORATORY. **IN-SPIRE™**. Disponível em: <https://in-spire.pnnl.gov/index.stm>. Acesso em: 4 maio. 2019.

PAN, X.; YAN, E.; CUI, M.; HUA, W. Examining the usage, citation, and diffusion patterns of bibliometric mapping software: A comparative study of three tools. **Journal of Informetrics**, v. 12, n. 2, p. 481–493, 2018.

PERSSON, O.; DANELL, R.; SCHNEIDER, J. W. How to use Bibexcel for various types of bibliometric analysis. **Celebrating scholarly communication studies: A festschrift for Olle Persson at his 60th birthday**, v. 5, p. 9–24, 2009.

PORTER, A. L.; CUNNINGHAM, S. W. Tech mining. **Competitive Intelligence Magazine**, v. 8, n. 1, p. 30–36, 2005.

SCI2 TEAM. **Science of Science (Sci2) Tool**. 2009. Disponível em: <https://sci2.cns.iu.edu/user/index.php>. Acesso em: 4 maio. 2019.

SCIMAGO. **SCImago Journal & Country Rank**. 2019. Disponível em: <https://www.scimagojr.com/aboutus.php>. Acesso em: 4 maio. 2019.

THELWALL, M. Bibliometrics to webometrics. **Journal of Information Science**, v. 34, n. 4, p. 605-621, 2008.

THOMPSON, D. F.; WALKER, C. K. A descriptive and historical review of bibliometrics with applications to medical sciences. **Pharmacotherapy: The Journal of Human Pharmacology and Drug Therapy**, v. 35, n. 6, p. 551-559, 2015.

VAN ECK, N.; WALTMAN, L. Software survey: VOSviewer, a computer program for bibliometric mapping. **Scientometrics**, v. 84, n. 2, p. 523–538, 2009.

VAN ECK, N. J.; WALTMAN, L. VOS: A new method for visualizing similarities between objects. In: DECKER, R.; LENZ, H. (Eds.). **Advances in data analysis**. Springer: Berlim, 2007. p. 299-306.

VAN ECK, N. J.; WALTMAN, L. Visualizing bibliometric networks. In: ROUSSEAU, R.; WOLFRAM, L. (Eds.). **Measuring scholarly impact**. Springer: Cham, 2014. p. 285-320.

VAN ECK, N. J.; WALTMAN, L. **VOSviewer Manual**, 2018.

ZUPIC, I.; ČATER, T. Bibliometric methods in management and organization. **Organizational Research Methods**, v. 18, n. 3, p. 429-472, 2015.