

A PRODUÇÃO CIENTÍFICA SOBRE CIÊNCIA DA INFORMAÇÃO E HUMANIDADES DIGITAIS INDEXADA NAS BASES DE DADOS DIMENSIONS, SCOPUS E WEB OF SCIENCE

*Fabiane Führ**

Programa de Pós-Graduação em Ciência da Informação (PGCin), Universidade Federal de Santa Catarina (UFSC).

*Edgar Bisset Alvarez***

Programa de Pós-Graduação em Ciência da Informação (PGCin), Universidade Federal de Santa Catarina (UFSC).

*Paula Carina de Araújo****

Programa de Pós-Graduação em Gestão da Informação (PPGI), Universidade Federal do Paraná (UFPR).

Resumo: Analisa os artigos científicos sobre Ciência da Informação e Humanidades Digitais indexados nas bases de dados Dimensions, Scopus e Web of Science sob uma perspectiva bibliométrica. Desenvolve uma pesquisa exploratória e um estudo bibliométrico com 57 artigos publicados entre os anos de 2011 a 2020. A publicação em inglês é a que predomina. Os anos de 2015 e 2018 foram os mais produtivos. A autoria compartilhada está presente em 64,92% dos artigos. O periódico mais produtivo é o *Proceedings of the Association for Information Science and Technology (ASIS&T)*. Quando analisados os resumos observa-se que as palavras recorrentes passam a ser ciência, humanidades, tecnologia, comunidade, humanidades digitais, biblioteca, entre outras que constam na rede de palavras. Conclui que essa temática permite diversas pesquisas e interações entre a Ciência da Informação e as Humanidades Digitais e que este corpus pode ser explorado em outras investigações, como o acoplamento bibliográfico, a análise de cocitação, a análise do domínio, entre outros.

Palavras-chave: Ciência da Informação; humanidades digitais; bibliometria, estudos métricos; interdisciplinaridade.

Título: PRODUCCIÓN CIENTÍFICA SOBRE CIENCIAS DE LA INFORMACIÓN Y HUMANIDADES DIGITALES INDEXADA EN LAS BASES DE DATOS DIMENSIONS, SCOPUS Y WEB OF SCIENCE.

Resumen: Analiza artículos científicos sobre Ciencias de la Información y Humanidades Digitales indexados en las bases de datos Dimensions, Scopus y Web of Science bajo una perspectiva bibliométrica. La investigación es considerada como exploratoria, siendo realizado un estudio bibliométrico con 57 artículos publicados entre los años 2011 a 2020. La mayoría de las publicaciones recuperadas están en inglés. Los años 2015 y 2018 fueron los más productivos. La autoría colaborativa está presente en el 64,92% de los artículos. La revista más productiva es *Proceedings of the Association for Information Science and Technology (ASIS & T)*. Cuando se analizan los resúmenes, se observa que las palabras recurrentes pasan a ser ciencia, humanidad, tecnología, comunidad, humanidades digitales, biblioteca, entre otras que aparecen en la red de palabras. Se concluye que esta temática permite varias investigaciones e interacciones entre las Ciencias de la Información y las Humanidades Digitales y que este corpus puede ser explorado en otro tipo de investigaciones como el acoplamiento bibliográfico, el análisis de cocitación, el análisis de dominio, entre otros.

Palabras clave: Ciencia de la información; humanidades digitales; bibliometría; estudios métricos; interdisciplinariedad.

Title: THE SCIENTIFIC OUTPUT ABOUT INFORMATION SCIENCE AND DIGITAL HUMANITIES INDEXED ON DIMENSIONS, SCOPUS AND WEB OF SCIENCE DATABASES.

Abstract: It analyzes the scientific articles on Information Science and Digital Humanities indexed in the databases Dimensions, Scopus and Web of Science from a bibliometric perspective. It develops an exploratory and bibliometric study of 57 articles published between the years 2011 to 2020. The majority of publications are in English. The years 2015 and 2018 were the most productive. Shared authorship with 64,92% is the predominant one. The most productive journal is the *Proceedings of the Association for Information Science and Technology* which publishes on information science and information technology. When analyzing the abstract it is observed that the recurring words become science, humanity, technology, community, digital humanities, library among others that are in the word network. It concludes that this subject allows several investigations and interactions between Information Science and Digital Humanities. Furthermore, the corpus can be explored in other studies such as bibliographic coupling, cocitation analysis, domain analysis, among others.

Keywords: Information science; digital humanities; bibliometric; metric studies; interdisciplinarity.

* fabiane.fuhr@ufpr.br

** edgar.bisset@ufsc.br

*** paulacarina@ufpr.br

Recibido: 13-05-2021; 2ª versión: 31-07-2021; aceptado: 15-09-2021.

Copyright: © 2021 Servicio de Publicaciones de la Universidad de Murcia (Spain). Este es un artículo de acceso abierto distribuido bajo los términos de la licencia Creative Commons Reconocimiento 4.0 Internacional (CC BY 4.0).

1 INTRODUÇÃO

A área de Humanidades, compreendida pelas disciplinas de História, Literatura, Artes, Filosofia, Educação, Ciências Sociais, Música, entre outras, tem sido transformada pelo uso das tecnologias da informação. Dessa forma, ao uso das tecnologias da informação para o desenvolvimento de pesquisa nas áreas de humanidades e ciências sociais deu-se o nome de Humanidades Digitais (HD).

Uma visão mais profunda de um domínio é proporcionada a partir da análise da sua produção científica e, nesse contexto é possível evidenciar a importância dos estudos métricos da informação, que permitem reconhecer os indicadores e relações, por exemplo, no que diz respeito a determinada literatura científica. A combinação dos estudos métricos com outras abordagens possibilita o conhecimento teórico, metodológico e epistemológico de um domínio (Hjørland, 2002; Oliveira, 2018; Grácio, 2020).

Marcelo Nogueira de Siqueira e Daniel Flores (2019) percebem a ciência da informação como uma área de interseção com as humanidades digitais. Para os autores as humanidades digitais compreendem “um território transdisciplinar que objetiva a divulgação, circulação, valorização e preservação do conhecimento e da pesquisa” (Siqueira y Flores, 2019, p. 85). Apresentam ainda como outro objetivo diretamente relacionado à ciência da informação, o “livre acesso aos seus dados e metadados através de ferramentas e possibilidades que a tecnologia e o ambiente digital proporcionam [...]” (Siqueira y Flores, 2019, p. 85).

Na busca por aprofundar os conhecimentos sobre os domínios das áreas de Ciência da Informação (CI) e HD, o objetivo geral deste trabalho é analisar os artigos científicos publicados sobre ambas as áreas indexados nas bases de dados Dimensions, Web of Science e Scopus sob uma perspectiva bibliométrica. Para alcançar o objetivo geral, definiram-se os seguintes objetivos específicos: a) identificar a elite de pesquisadores da área de Humanidades Digitais e Ciência da Informação; b) aplicar a Lei de Bradford para identificar os periódicos mais produtivos; c) apresentar os indicadores de produção relacionados ao corpus da pesquisa e d) demonstrar a frequência da coocorrência de termos atribuídos pelos autores como palavras-chave e dos termos contidos nos abstracts.

A pesquisa se inicia com a apresentação do tema e objetivos nesta introdução, na seção seguinte é apresentado o referencial teórico. Posteriormente, apresenta-se a metodologia, seguida da apresentação e discussão dos resultados e, por fim são apresentadas as considerações finais e as referências.

2 A CIÊNCIA DA INFORMAÇÃO E AS HUMANIDADES DIGITAIS

O surgimento da CI está atrelado ao desenvolvimento tecnológico, porém desde o princípio propõe-se a investigar as propriedades e o comportamento informacional, as forças que governam o fluxo da informação e os meios de processamento da informação, visando a acessibilidade e a usabilidade das ideias. Este domínio está preocupado com o corpo de conhecimentos relacionados à origem, a coleta, a organização, ao armazenamento, a recuperação, a interpretação, a transmissão, a transformação e a utilização da informação (Borko, 1968).

Desde sua gênese a CI apresentou características interdisciplinares e contou com diferentes profissionais como engenheiros, bibliotecários, químicos, linguistas, filósofos, psicólogos, matemáticos, cientistas da computação, homens de negócios e tantos outros profissionais ligados ou não a ciência (Saracevic, 1996). Essa característica interdisciplinar manteve-se ao longo do desenvolvimento da CI, uma vez que os diálogos e as interações com as outras disciplinas sempre estiveram presentes e têm se fortalecido desde a sua criação.

O mesmo tem acontecido com as HD, que podem ser compreendidas como um campo transdisciplinar que envolve pesquisas e práticas das Ciências Humanas e Sociais ao se relacionar com o ambiente digital, principalmente, no que se refere ao acesso à informação, aos dados e a produção do conhecimento (Andrade y Dal'Evedove, 2020).

Segundo Russell (2011) os objetivos da HD são a criação de bases de dados com recursos digitais relevantes para as Humanidades, que incluam a captura, a estruturação, a documentação, a preservação e a disseminação dos dados; desenvolvam metodologias que permitam gerar novos elementos derivados destes dados e gerem investigações e conhecimentos para incrementar a compreensão sobre as Humanidades.

Dessa forma, o que se observa é que há uma convergência entre as práticas realizadas pelas duas disciplinas, HD e CI; ambas têm seu foco e prática em informações registradas nos documentos (Robinson, Priego y Bawden, 2015; Koltay, 2015); muitos projetos dessas disciplinas estão voltados para pesquisas sobre patrimônio cultural, que incluem digitalização, preservação, desenvolvimento de coleções, catalogação, visualização e classificação (Robinson, Priego y Bawden, 2015; Sula, 2013); e elas se depararam com os desafios impostos pelas tecnologias, na qual “dados diferentes, advindos de bases e heterogêneas encontram por mediações tecnológicas possibilidades de produzir informações intercruzadas” (Pimenta, 2016, p. 22).

Além dos pontos anteriormente mencionados, que são objetos de estudos de várias pesquisas realizadas no âmbito da CI e também das HD, Robinson, Priego e Bawden (2015) observaram que há várias outras temáticas que podem ser de interesse para ambas as áreas, como os estudos referentes à busca e recuperação da informação; as bibliotecas e os arquivos digitais; os metadados e a descrição de recursos; as ontologias, classificações e taxonomias; os meios de publicação e divulgação científica; o Acesso Aberto; o Linked Data (dados vinculados); a gestão e a curadoria de coleções; os portais de periódicos e os repositórios; a bibliografia; a digitalização e seus processos; a preservação; a preservação digital; a interatividade e a experiência do usuário, interfaces e navegação; a herança cultural; a visualização da informação; o big data; a mineração de dados e a bibliometria.

Andrade e Dal'Evedove (2020) apontam que a CI normalmente discute como se apropriar da tecnologia e da inovação em benefício das pessoas, das organizações e para a sociedade, enquanto as HD surgem como um movimento que se apoia nas tecnologias computacionais para permanecer atual frente aos objetivos da sociedade da informação. Assim, como afirma Pimenta (2020), HD e CI não deveriam ser pensadas de modo separado, uma vez que ambas têm a informação como seu foco central, mas poderiam ser pensadas de forma complementar. É nesse contexto que esta pesquisa foi desenvolvida, ao buscar reconhecer a aproximação desses dois domínios na literatura científica internacional.

3 METODOLOGIA

Esta pesquisa se caracteriza como um estudo bibliométrico aplicado a uma disciplina científica, ou seja, aplicado à Ciência da Informação buscando identificar as inter-relações existentes entre a Ciência da Informação (CI) e as Humanidades Digitais (HD). Assim, a pesquisa desenvolvida neste artigo é do tipo exploratória quanto aos seus objetivos. Este estudo explora a produção científica relacionada à intersecção entre a Ciência da Informação e as Humanidades Digitais. Quanto aos procedimentos, desenvolve-se um estudo bibliométrico, pois descreve e identifica os seguintes indicadores de produção: ano de publicação, idioma, tipos de autoria, produtividade de autores e periódicos, analisa a colaboração entre os autores mais produtivos e analisa a coocorrência de termos a partir das palavras-chave e resumos dos artigos que compõem o corpus desta pesquisa.

Os Estudos Métricos da Informação (EMI) são o conjunto de conhecimentos relativos à avaliação da informação produzida, fundamentada na Sociologia da Ciência, na Ciência da Informação, na Matemática, na Estatística e na Computação. Os estudos podem ser de natureza teórico-conceitual, quando contribuem para o avanço do conhecimento da própria temática ao propor novos conhecimentos e indicadores, assim como novas reflexões e análises referentes à área. Também podem ser de natureza metodológica quando dão sustentação a trabalhos de caráter teórico da área na qual estão sendo aplicados (Oliveira y Grácio, 2016; Oliveira, 2018).

Glänzel (2003, p. 9) considera que a pesquisa bibliométrica é destinada a três grupos-alvo principais, que determinam os tópicos e as subáreas da Bibliometria contemporânea, a saber: **Bibliometria para profissionais da bibliometria (G1)**: busca desenvolver e debater a bibliometria como metodologia, isto é, está preocupada com o próprio desenvolvimento conceitual-teórico-metodológico. **Bibliometria aplicada às disciplinas científicas (G2)**: desenvolve pesquisas bibliométricas “aplicadas” e forma o maior e mais diversificado grupo de interesse na bibliometria. **Bibliometria para a política científica e gestão (G3)**: refere-se a avaliação da pesquisa com fins de orientar políticas científicas. Para Glänzel este grupo-alvo é o tópico mais importante da bibliometria contemporânea. Esta pesquisa tem aderência ao G2 de Glänzel, pois utiliza-se a bibliometria para reconhecer a relação entre dois domínios.

Curty e Delbianco (2020) descrevem em pesquisa recente os principais estudos métricos da informação, a saber: Almetria, Arquivometria, Bibliometria, Cibermetria, Cientometria, Informetria, Patentometria e Webometria. Esta pesquisa tem como foco a informação científica registrada e os seus indicadores de produção, por isso, apoiou-se na Bibliometria para chegar aos resultados pretendidos. Destaca-se que, assim como Vanti (2002), compreende-se como potencialidades da aplicação desses estudos, a identificação das tendências e o crescimento do conhecimento do domínio

aqui estudado, a possibilidade de prever a produtividade de autores individuais, organizações e países, medir o grau de colaboração entre autores e o crescimento de determinadas áreas e o surgimento de novos temas.

Com o objetivo de identificar a produção científica na área de Ciência da Informação e Humanidades Digitais definiu-se como estratégia de busca a expressão [("ciência da informação" OR "ciencia de la información" OR "information science") AND ("humanidade digital" OR "humanidades digitais" OR "humanidades digitales" OR "digital humanity" OR "digital humanities")]. A estratégia combinou os termos “Ciência da Informação” e “Humanidades Digitais” nos idiomas Português, Espanhol e Inglês e foi aplicada nas bases de dados interdisciplinares Dimensions, Scopus e Web of Science (WoS).

Para esta pesquisa utilizamos a Dimensions Free, uma versão gratuita e aberta da base, formada colaborativamente e lançada em 2018. São também mantidos outros três produtos pagos, Dimension Analytics, Dimension Profile e Dimensions on Google BigQuery. A base de dados Dimensions é uma base que integra publicações, citações, concessões, patentes, ensaios clínicos, conjuntos de dados e documentos de política que podem ser acessados de forma gratuita. A *Digital Science* e as empresas ReadCube, Altmetric, Figshare, Symplectic, Digital Science Consultancy e ÜberResearch trabalham em colaboração com a comunidade científica no desenvolvimento e melhoria contínua da base (Dimensions, 2021). Na Dimensions a estratégia de busca foi aplicada nos campos título e no resumo.

A Web of Science é uma base de dados multidisciplinar e pioneira na área, criada em 1997, que reúne diferente índices de citação. Atualmente, é mantida pela Clarivate Analytics. Já a Scopus é a base de dados criada em 2004 e mantida pela Elsevier e composta por um índice de citação. As duas fontes são pagas e o acesso a elas foi possível por meio do Portal de Periódicos Capes. Nas bases de dados Web of Science e Scopus a estratégia de busca foi aplicada nos campos título, resumo e palavras-chave.

A Dimensions foi escolhida como fonte para a coleta dos dados deste estudo com base em pesquisas que já analisaram sua cobertura e funcionalidade e afirmaram que ela é uma alternativa para estudos métricos da informação, fazendo frente à bases como Web of Science e Scopus (Orduña-Malea, Delgado-López-Cózar, 2018; Harzing, 2019). Todavia, as bases Web of Science e Scopus também foram utilizadas como fonte de informação para esta pesquisa, pois são consideradas as principais bases de dados multidisciplinares e que tem como foco métricas de informação científica (Harzing y Alakangas, 2016; Moed, 2017; Sugimoto y Larivière, 2018).

Dimension é uma ferramenta é completamente transparente em termos de dados, disponibiliza estatística completa do total de documentos, oferece pesquisa por palavras-chave ou por resumo. A pesquisa por palavras-chave pode ser realizada no texto completo ou no resumo e título. Após a conclusão da pesquisa, a ferramenta oferece uma série de filtros, tais como: ano de publicação, pesquisador, área do conhecimento, tipo de publicação, fonte de publicação, lista de periódicos e documentos em acesso aberto (Orduña-Malea y Delgado-López-Cózar, 2018). A estratégia de busca foi aplicada no dia 02 de janeiro de 2021. Em todas as bases aplicou-se o filtro de publicação, limitando a busca aos artigos. Quanto ao período de publicação, foram encontrados e coletados artigos publicados entre os anos de 2011 a 2020. Na próxima seção serão apresentados os resultados e analisados os indicadores de produção dos artigos indexados nas bases de dados Dimensions, Scopus e WoS.

4 ANÁLISE DA PRODUÇÃO CIENTÍFICA ENTRE CIÊNCIA DA INFORMAÇÃO E HUMANIDADES DIGITAIS

O levantamento retornou 102 resultados, apresentados no Quadro 1. Utilizou-se o gerenciador bibliográfico Zotero para a coleta, o armazenamento e a organização dos artigos recuperados no levantamento bibliográfico.

BASES DE DADOS MULTIDISCIPLINARES	
Base de dados	Artigos
Dimensions	34
Scopus	35
Web of Science	33
TOTAL	102

Quadro 1. Pesquisa realizada em bases de dados. Fonte: Dados da pesquisa (2021).

Após a organização e padronização do metadados foram excluídos um total de catorze (14) documentos e a seleção se deu por idioma e tipo de documentos. Destes, foram excluídos dois (2) artigos em húngaro, dois (2) artigos em

italiano, dois (2) artigos em polonês, quatro (4) artigos em russo e um (1) artigos em ucraniano, dois (2) resumos e um (1) editorial. Estes artigos foram excluídos, devido às barreiras linguísticas, os resumos e o editorial foram excluídos por não se enquadrarem no foco e escopo desta pesquisa. Portanto, após a organização, padronização e exclusão desses documentos, o corpus da pesquisa para o estudo bibliométrico obteve um total de 88 artigos.

BASES DE DADOS MULTIDISCIPLINARES	
Base de dados	Artigos
Dimensions	26
Scopus	34
Web of Science	28
TOTAL	88

Quadro II. Resultado após padronização. Fonte: Dados da pesquisa (2021).

Todavía, uma nova análise mostrou que alguns artigos foram indexados em duas, ou nas três bases, conforme a representação gráfica apresentada no Gráfico 1. Dessa forma, apesar do resultado da padronização ser de 88 artigos, após a eliminação das repetições restaram 57 artigos para o estudo.

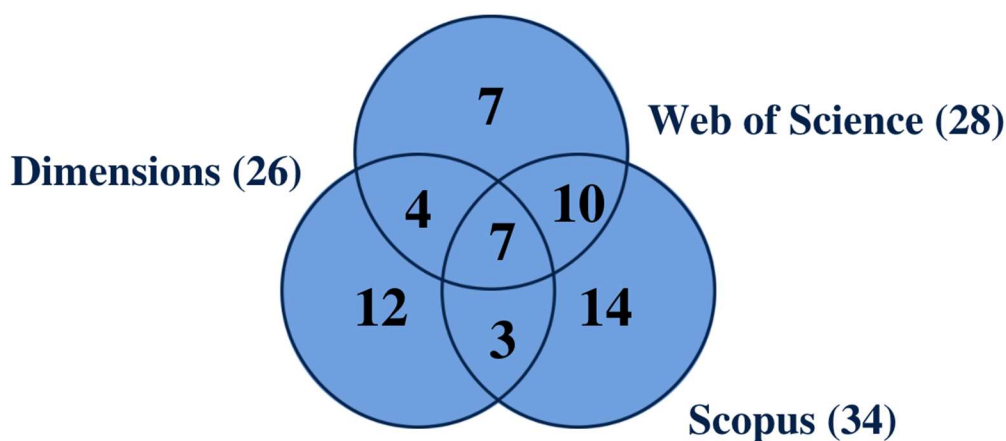


Gráfico 1. Gráfico de intersecção entre as bases de dados. Fonte: Dados da pesquisa (2021).

Depois de definido o corpus para análise da pesquisa, os dados foram exportados do software Zotero, em formato csv, para geração de tabelas, gráficos e outras análises utilizando o software Excel. Também foi gerado um arquivo em formato ris, para a criação da rede de coautoria, para a rede de coocorrência de termos extraídos das palavras-chave e para a rede de coocorrência de termos extraídos dos abstracts por meio da utilização do software VOSviewer que são apresentados ao longo desta pesquisa.

A análise demonstra que a temática vem ganhando força ao longo dos anos. No gráfico 2 é possível observar que nos 3 primeiros anos há poucos estudos publicados, porém já em 2015 ocorre um aumento significativo no número de publicações (11 artigos), depois há uma leve queda nos 2 anos seguintes, atingindo novamente 11 artigos publicados em 2018 e mantendo-se estável nos 2 anos seguintes.

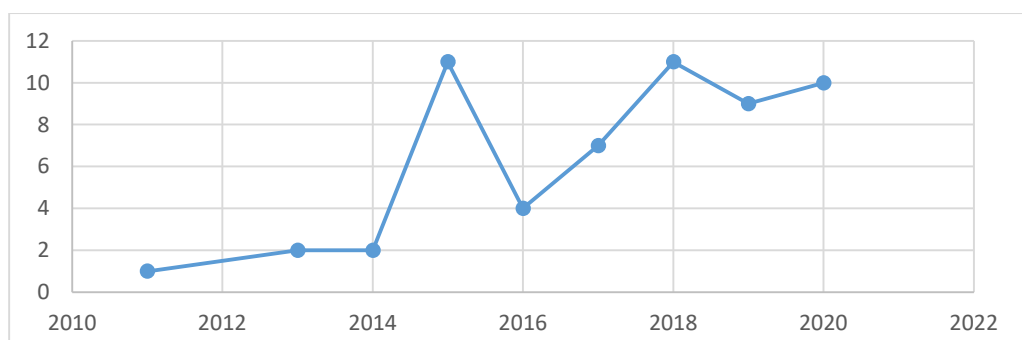


Gráfico 2. Artigos por ano de publicação. Fonte: Dados da pesquisa (2021).

Este aumento pode ser justificado, pois “as Humanidades Digitais passam a integrar a Biblioteconomia e a Ciência da Informação na medida em que produtos e serviços provenientes de sistemas de recuperação de informação convergem com o ambiente digital” (Andrade y Dal’Evedove, 2020, p. 441). Todavia, os limites entre um e outro permanecem imprecisos e necessitam de maiores aprofundamentos.

Com relação ao idioma dos artigos, observa-se que o inglês é o idioma predominante com 37 artigos. O português é o segundo idioma mais frequente com 11 artigos publicados, seguido do alemão (5) e do espanhol (4).

O tipo de autoria foi outro indicador analisado, conforme demonstrado na Tabela I. Pode-se observar que a autoria compartilhada é a adotada na maioria dos artigos com 64,92% e a autoria única corresponde a 35,08% dos artigos.

Tipo de autoria	Autores	Artigos
Autoria única	1 autor	20
Autoria compartilhada	2 autores	20
	3 autores	11
	4 autores	2
	5 autores	2
	6 autores	2
TOTAL		57

Tabela I. Tipo de autoria dos artigos do corpus da pesquisa. Fonte: Dados da pesquisa (2021).

A análise de autoria revela que os 57 artigos do corpus apresentam 123 autores. A Tabela II apresenta os 6 autores mais produtivos, suas respectivas instituições, países e total de artigos. Esta tabela apresenta os autores que tem dois artigos ou mais no corpus analisado nesta pesquisa, ou seja, dos 6 autores, 3 publicaram 3 artigos cada um e 3, publicaram 2 artigos cada um. Apenas 2 dos autores mais produtivos possuem algum vínculo entre si. Os autores Crista Womser-Hacker e Ben Heuwing apresentam o mesmo vínculo institucional, ambos atuam na Universidade de Hildesheim, na Alemanha. Há predominância de autores vinculados a instituições europeias entre os 6 mais produtivos.

Autores	Instituição	País	Artigos
Pimenta, Ricardo Medeiros	Instituto Brasileiro de Informação em Ciência e Tecnologia	Brasil	3
Purves, Ross S.	Universidade de Zurique	Suíça	3
Womser-Hacker, Christa	Universidade de Hildesheim	Alemanha	3
Engerer, Volkmar P.	Universidade de Copenhage	Dinamarca	2
Hernández Quintana, Ania Rosa	Universidade de Havana	Cuba	2
Heuwing, Ben	Universidade de Hildesheim	Alemanha	2

Tabela II. Autores mais produtivos. Fonte: Dados da pesquisa (2021).

Quanto ao tipo de autoria praticada por estes autores, é possível observar que Ricardo Medeiros Pimenta possui 2 artigos de autoria individual e 1 de autoria coletiva (1 artigo publicado com 2 autores). O autor Ross S. Purves publicou todos os artigos em coautoria, destes 1 artigo com 2 autores e 2 com 3 autores. Constatou-se que há relação de coautoria entre dois dos autores mais produtivos, Crista Womser-Hacker e Bem Heuwing (2 artigos). Há predominância de artigos publicados em coautoria entre os autores mais produtivos.

Para visualizar as relações entre os autores, foi criada uma rede de coautoria obtida por meio do software VOSviewer, na qual os autores são representados pelas esferas e as linhas representam as ligações existentes entre os autores. Quanto maior a esfera, maior é o número de publicações desse autor no corpus da pesquisa e quanto mais espessa a linha, mais forte é a ligação entre os autores. Ao analisar a Figura 1, é possível observar a formação dos clusters conforme o tipo de autoria identificada na Tabela 1.



Figura 1. Rede de coautoria.

A Figura 1 foi obtida por meio do método de contagem completa com restrição mínima de 1 (um) autor por documento. Foram identificados 113 autores, que formaram 49 clusters. Há 3 clusters (verde, azul e vermelho) com 6 itens. Nos clusters, azul e vermelho, cada um dos itens do cluster apresenta 5 ligações, o que mostra apenas uma ligação com os demais autores do próprio documento analisado. O cluster verde também apresenta 6 itens, porém o que se observa é que o autor Purves, Ross S. é autor de 3 documentos que dão origem a 5 ligações que compõem este cluster, como é possível observar na Figura 2.

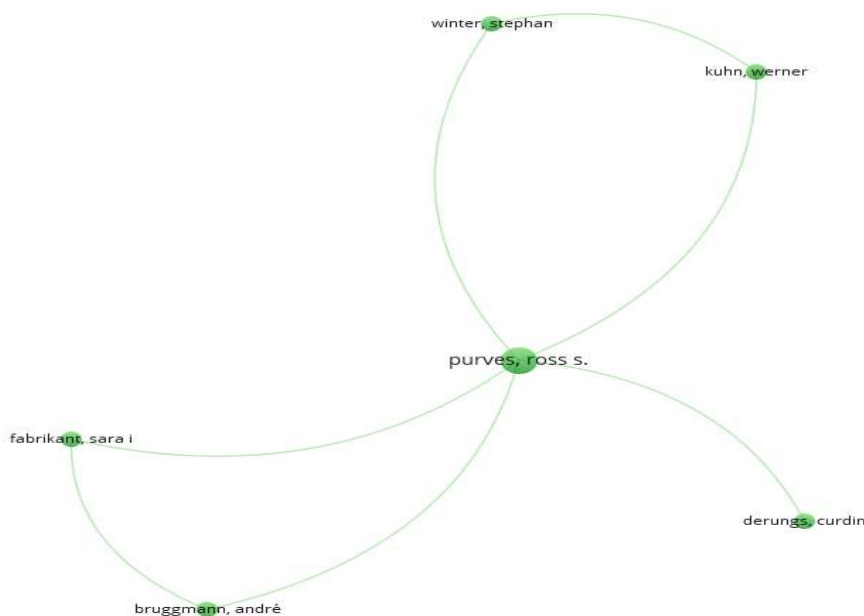


Figura 2. Cluster de Ross S. Purves. Fonte: Dados da pesquisa (2021).

Foram identificados 3 clusters com 5 itens (azul turquesa, lilás e amarelo). Os clusters azul turquesa e lilás são formados, ambos por apenas 1 documento e cada um dos clusters apresenta 4 links. O cluster amarelo, cuja autora central é Womser-Hacker, Christa apresenta 3 documentos, os quais formam 5 ligações (Figura 3). Também foram identificados 1 cluster com 4 itens, 8 clusters com 3 itens, 17 clusters com 2 itens e 16 clusters com 1 item.



Figura 3. Cluster de Crista Womser-Hacker. Fonte: Dados da pesquisa (2021).

Outra análise realizada no corpus da pesquisa, foi com relação a avaliação do conjunto de periódicos científicos. Para essa avaliação, utilizou-se a Lei de Bradford, segunda lei da bibliometria, que permite avaliar a produtividade desses periódicos. Assim, ao ordenar um conjunto de periódicos em ordem de produtividade decrescente relevante a um determinado assunto, será possível observar 3 zonas, cada uma com 1/3 do número total de artigos relevantes. Dessa forma, a primeira zona conterá um pequeno número de periódicos altamente produtivos, a segunda conterá um número maior de periódicos menos produtivos e a terceira zona incluirá ainda mais periódicos, com menor produtividade (Araújo, 2006).

Na Tabela III é possível observar as zonas de produtividade obtidas ao aplicar a Lei de Bradford ao conjunto de periódicos do corpus da pesquisa.

Zonas	Quantidade de periódicos	Artigos
Zona 1 (periódicos mais produtivos)	5 periódicos	19
Zona 2	15 periódicos	19
Zona 3 (periódicos menos produtivos)	19 periódicos	19
TOTAL	39 periódicos	57

Tabela III. Produtividade de periódicos do corpus da pesquisa. Fonte: Dados da pesquisa (2021).

Percebe-se que o Zona 1 apresenta um grupo reduzido de periódicos muito produtivos, o que confirma o enunciado da Lei de Bradford. Estes periódicos foram reunidos na Tabela IV, na qual foram incluídas informações como nome do periódico, ISSN, número de artigos, SCImago Journal Rank (SJR)¹, Fator de Impacto do Journal Citation Reports (JCR)² e país de origem da publicação.

Título do periódico	ISSN	Artigos	SJR ³ (Scopus)	Fator de Impacto ⁴ (JCR/WoS)	País
Proceedings of the Association for Information	0044-7870	8	0.27/2019	*	EUA

Título do periódico	ISSN	Artigos	SJR ³ (Scopus)	Fator de Impacto ⁴ (JCR/WoS)	País
Science and Technology (ASIS&T)					
Journal of Documentation	0022-0418	4	0.89/2019	1.819/2020	Reino Unido
Information - Wissenschaft und Praxis	1619-4292	3	0.13/2019	*	Alemanha
Digital Library Perspectives	2059-5816	2	0.34/2019	*	Reino Unido
Informação & Sociedade: Estudos	0104-0146	2	0.26/2019	0.311/2020	Brasil

* Não foi identificada informação para este periódico.

Tabela IV. Periódicos mais produtivos (zona 1). Fonte: Dados da pesquisa (2021).

O periódico mais produtivo é o *Proceedings of the Association for Information Science and Technology (ASIS&T)* (0044-7870), com periodicidade anual e que deu continuidade ao *Proceedings of the ASIST Meeting* encerrado em 2014. Este periódico publica sobre as temáticas de ciência da informação, tecnologia da informação, metodologia da informação, entre outros. Os 8 artigos do periódico mais produtivo foram publicados na seguinte ordem cronológica: 1 em 2014, 3 em 2015, 1 em 2016, 1 em 2017 e 3 em 2018. Dessa forma, é possível observar a frequência e a relevância dessa temática para a publicação.

O *Journal of Documentation* (0022-0418) é uma publicação bimestral, que publica resultados de todas as disciplinas vinculadas a informação. Quanto aos artigos publicados, 1 artigo é de 2011, 1 artigo de 2016 e 2 artigos foram publicados em 2020. O periódico *Information - Wissenschaft und Praxis* (1619-4292), bimestral, é um periódico profissional que publica há mais de 65 anos informações sobre ciência da informação na Alemanha. Os 3 artigos do corpus foram publicados em 2015.

O periódico *Digital Library Perspectives* (2059-5816), é uma publicação trimestral, que publica pesquisas relacionadas à curadoria, preservação e arquivamento digital, objetos digitais, bibliotecas e repositórios digitais. Os 2 artigos do corpus foram publicados em 2018. O periódico *Informação & Sociedade: Estudos* (0104-0146) é uma revista trimestral que divulga trabalhos em ciência da informação, biblioteconomia e áreas afins. Os artigos do corpus (2) foram publicados 1 em 2018 e outro em 2020. Constata-se que a presença da temática das HD na CI nos periódicos mais produtivos do corpus desta pesquisa se deu a partir de 2014, ou seja, trata-se de um tema recente.

Para a análise de coocorrência de termos observou-se que 9 artigos do *corpus* não apresentavam palavras-chave, logo não foram contabilizados para a análise de coocorrências de palavras-chave. A segunda ação adotada foi utilizar as palavras-chave em inglês (keywords) adotadas pelos autores ou traduzir as palavra-chave para os artigos publicados em outros idiomas que não continham palavras-chave em inglês.

As palavras-chave utilizadas pelos autores também foram avaliadas (Figura 4). Nos 57 artigos do corpus foram identificadas 245 palavras-chave que foram consideradas para a análise de coocorrência de palavras-chave. Ao aplicar as palavras-chave no software VOSviewer, mantendo o método de contagem completo e a restrição mínima de coocorrência de palavra-chave em 1, identificou-se 182 palavras-chaves, destas foram excluídas as palavras-chave *digital humanities* e *information science*, pois são as palavras que deram origem a pesquisa e por isso seriam as mais recorrentes. Das 180 palavras-chave restantes foram identificadas 48 que estão conectadas entre si e que são apresentadas na Figura 4.

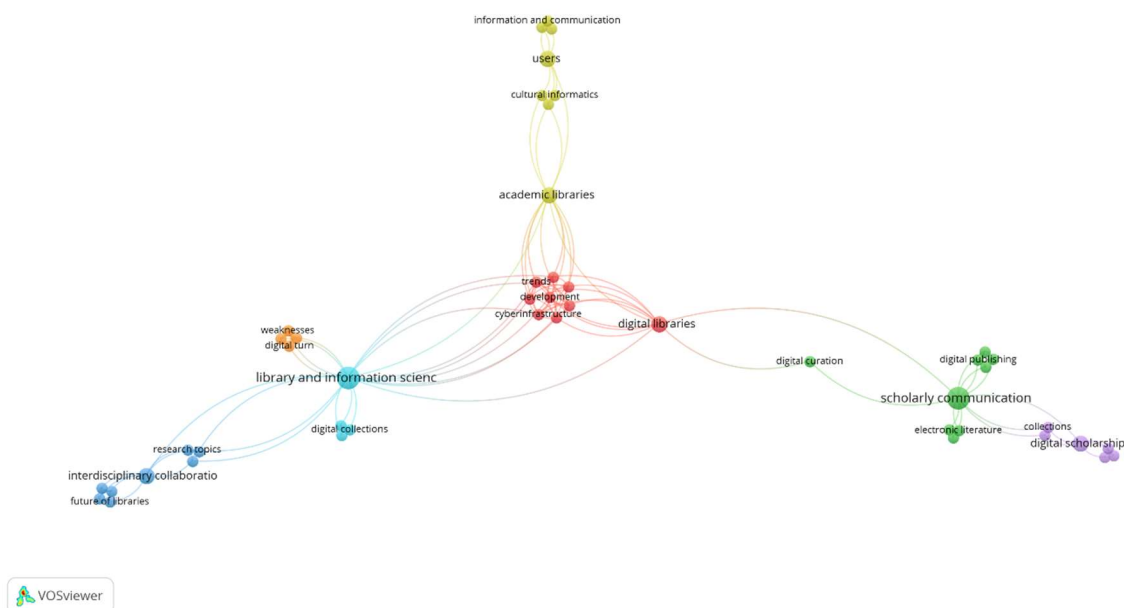


Figura 4. Coccorrência de termos extraídos das palavras-chave. Fonte: Dados da pesquisa (2021).

Entre os 48 itens foram identificados 7 clusters e 138 ligações entre os termos. O cluster 1 é formado por 9 itens e está apresentado pela cor vermelha. O item de mais relevância neste cluster é *digital libraries* com 2 ocorrências e 12 links, esse também é o termo que liga o cluster vermelho ao cluster amarelo e verde. O segundo cluster também apresenta 9 itens e este representado pela cor verde. O termo *scholarly communication* apresenta 4 ocorrências e 12 links.

O cluster 3 possui 8 itens e está representado pela cor azul. *Interdisciplinary collaboration* é o termo mais relevante deste cluster com 2 ocorrências e 8 ligações. No cluster 4, representado pela cor amarela os termos *academic library* e *users* apresentam 2 ocorrências, todavia o primeiro apresenta 14 ligações, enquanto o segundo apenas 7.

O quinto cluster está representado pela cor lilás, possui 6 itens e *digital scholarship* é o item com maior relevância com 2 ocorrências e 6 links. O cluster 6, representado pela cor azul turquesa, apresenta o termo *library and information science* com 4 ocorrências e 21 ligações. O sétimo cluster representado pela cor laranja apresenta 4 itens, cada um deles com uma ocorrência e 4 ligação cada.

A análise dos cluster e da relação entre os termos evidencia o caráter interdisciplinar da relação entre humanidades digitais e a ciência da informação, pois os termos *interdisciplinarity* e *interdisciplinary collaboration* estavam em destaque. A ocorrência de termo como *digital curation*, *cyberinfrastructure*, *digital libraries*, *digital scholarship* e *digital collections* demonstram os principais temas das pesquisas relacionadas aos dois domínios.

Além da análise da coocorrência de palavras-chave, fez-se uma análise com base no texto dos abstracts que compõem o corpus da pesquisa. Para isso, selecionou-se apenas os resumos dos artigos, o método de contagem adotado foi o método completo e o número mínimo de ocorrência de um termo foi parametrizado em 4 ocorrências. Dos 1562 termos, 107 atenderam ao limite estabelecido. O VOSviewer está parametrizado para selecionar 60% dos termos mais relevantes, dessa forma o sistema calculou que o número de termos selecionados será de 64 termos.

Observou-se que alguns resumos apresentavam os abreviaturas, dessa forma criou-se um tesauro para aplicar no VOSviewer.

Abreviatura	Termo
dh	digital humanities
dl	digital library
lis	library and information science
dh community	digital humanities community
dh research	digital humanities research

Quadro 3. Tesauros aplicados no VOSviewer. Fonte: Os autores (2021).

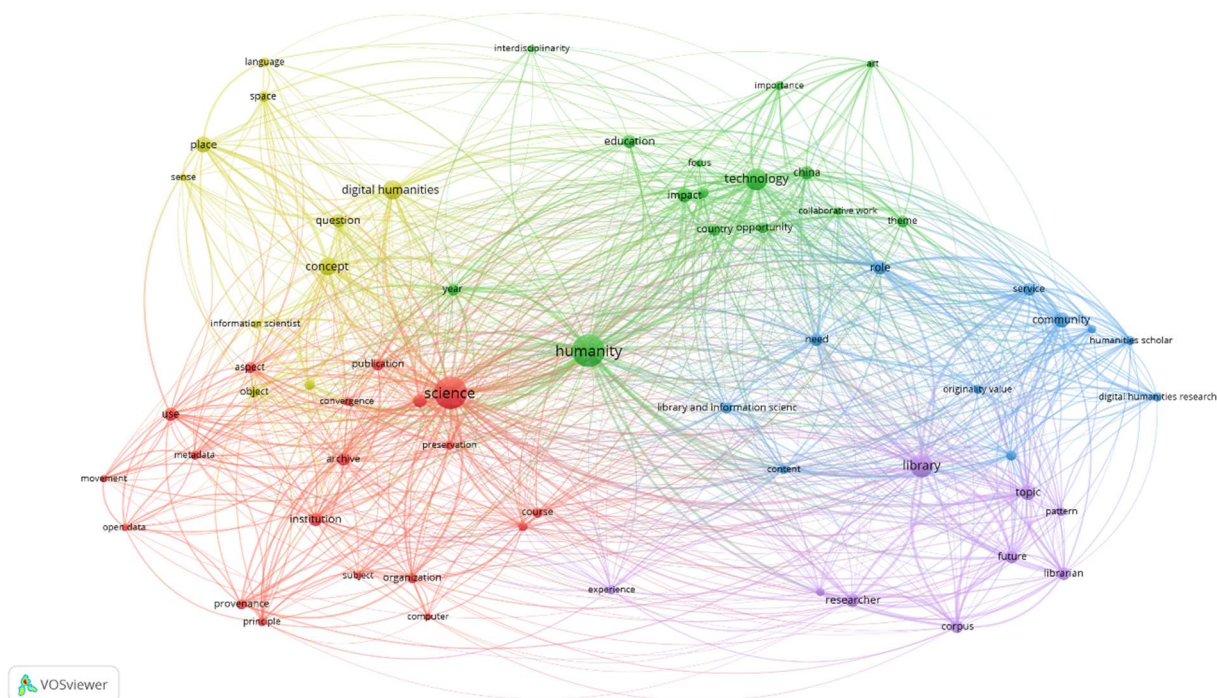


Figura 5. Coocorrência de termos extraídos dos abstracts. Fonte: Dados da pesquisa (2021).

O método de contagem adotado foi o método completo e o número mínimo de ocorrência de um termo foi parametrizado em 4 ocorrências. Dos 1562 termos, 107 atenderam ao limite estabelecido. O VOSviewer está parametrizado para selecionar 60% dos termos mais relevantes, dessa forma o sistema calculou que o número de termos selecionados será de 64 termos.

Estes 64 itens formaram 5 clusters, com 1005 conexões e 5728 ligações de força ente os termos. O primeiro cluster é formado por 19 itens, está representado na figura 6 pela cor vermelha. O termo *science* é o termo que apresenta mais ocorrências (67) neste cluster. O cluster 2 é formado por 15 itens e está representado na figura 6 pela cor verde, os termos de maior relevância neste cluster são *humanity* com 66 ocorrências e *technology* com 31 ocorrências. A força da ligação entre os termos é 51. Entre os termos *science* e *humanity* a força de ligação entre os termos é de 95.

O cluster 3 é formado por 11 itens e está representado pela cor azul. Os termos mais relevantes neste cluster são *community* com 15 ocorrências e *role* com 13 ocorrências. A força de ligação entre estes termos é de 25. O cluster 4 é formado por 10 itens e é representado pela cor amarela. Neste cluster os termos mais relevantes são *digital humanities* com 24 ocorrências e *concept* com 20 ocorrências. A força de ligação entre estes termos é 17. O cluster 5 é formado por 9 itens e representado pela cor lilás. O termo mais relevante neste cluster é *library* com 33 ocorrências e *topic* com 14 ocorrências. A força de ligação entre os termos é de 42.

Com base nos resumos e observação das figuras 4, 5 e 6 algumas informações ficam ainda mais evidentes. Os estudos apresentam pesquisas interdisciplinares sob diferentes perspectivas, tais como: relação entre HD e CI, impacto tecnológico, ações culturais, entre outras. Muitas pesquisas buscam compreender a relação entre a CI e as HD, principalmente sobre como a CI pode contribuir com estudos em humanidades, seja no desenvolvimento da estrutura digital, padrões de metadados, materiais digitalizados, objetos digitais genuínos, análise da informação geográfica (GIScience), mídias digitais, entre outras. A preocupação com a capacitação dos pesquisadores ligados às HD, o desenvolvimento de pesquisas colaborativas, o *crowdsourcing* e o trabalho voluntário também foram questão abordadas nos estudos do corpus desta pesquisa. A presença dos diferentes termos nas figuras 4 e 5 evidenciam o vasto campo de pesquisa das HD, reafirmando a importância desse estudo.

A temática humanidades digitais tem ganhado cada vez mais relevância no cenário nacional e internacional. Pode-se afirmar isso, com base no corpus da pesquisa, no qual se observa que 80,7% das publicações são nos idiomas inglês, alemão e espanhol. Todavia, a publicação em língua portuguesa merece destaque, pois 19,3% dos artigos que compõem a pesquisa foram publicados em português.

Observa-se que o número de publicações teve um pico em 2015 com 11 publicações. No ano seguinte, teve uma queda brusca, apresentando apenas 4 publicações para o ano de 2016, nos anos seguintes as publicações sobre a temática voltaram a crescer tendo seu novo ápice em 2018 (11 publicações) e mantendo-se estável nos anos seguintes. Acredita-se que há uma tendência de constância e até crescimento da quantidade de publicações se considerada a importância da temática para o domínio em questão.

A partir da análise bibliométrica aqui apresentada, foi possível identificar os principais periódicos que tem publicado sobre o tema e que devem estar no radar dos estudiosos do tema, entre eles se destaca o *Proceedings of the Association for Information Science and Technology*, que publica os anais de um dos principais eventos científicos da Ciência da Informação. No âmbito nacional, destaca-se o periódico *Informação & Sociedade*. Da mesma forma, reconhecem-se os autores seminais para a pesquisa que relaciona CI e HD, com destaque para autores da Europa, América Central e do Sul.

5 CONSIDERAÇÕES FINAIS

Neste estudo, constatou-se que as pesquisas sobre HD permitem uma série de análises a ela relacionadas, uma vez que o leque das humanidades é amplo, assim como o é, a CI, a junção de ambas as áreas leva a uma infinidade de novas possibilidades que podem ser exploradas por pesquisadores das duas áreas. Este estudo evidencia que as áreas são relacionadas e, por vezes, complementares e, também, demonstra que pesquisadores nacionais e internacionais tem se voltado para a temática de forma interdisciplinar.

A escolha de um número específico de bases de dados pode ser considerado uma limitação da pesquisa. Sugere-se novas pesquisas também com artigos indexados em outras bases de dados multidisciplinares como SciELO e Redalyc, o que pode aumentar o panorama de publicações em português e espanhol sobre HD e CI nos países latino-americanos. Além disso, também é possível considerar a replicação da pesquisa em bases de dados específicas da área de Ciências Sociais Aplicadas, que englobaria a CI e as HD.

Observa-se ainda que o corpus dessa pesquisa oferece potencial para o desenvolvimento de outros estudos métricos, como o acoplamento bibliográfico, a análise de cocitação e outras análises relacionais de citação. Some-se a esses estudos ainda a possibilidade de desenvolver uma análise de domínio e acrescentar maior profundidade ao conhecimento das questões epistemológicas, teóricas e metodológicas que envolvem o domínio.

NOTAS

¹ O SCImago Journal Rank (SJR) é um indicador que tem como objetivo mostrar a visibilidade dos periódicos indexados na Scopus desde 1996. (Scimago, 2021).

² O Journal Citation Reports (JCR) é um indicador que avalia e compara periódicos usando dados de citação obtidos por publicações técnicas e acadêmicas indexadas na Web of Science. (Portal..., 2021).

³ <<https://www.scimagojr.com/index.php>>

⁴ <<https://jcr-clarivate.ez22.periodicos.capes.gov.br/>>

6 REFERÊNCIAS

- ARAÚJO, C.A.Á. Bibliometria: evolução histórica e questões atuais. *Em Questão*, 2006, vol. 12, nº 1, p. 11-32. Disponível em: <<https://seer.ufrgs.br/EmQuestao/article/view/16>> [Acesso em: 21 abr. 2021]
- ANDRADE, L.M. de y DAL'EVEDOVE, P.R. Humanidades digitais na ciência da informação brasileira: análise da produção científica. *Revista Ibero-Americana de Ciência da Informação*, 2020, vol. 13, nº 1, p. 439-451. Disponível em: doi: 10.26512/rici.v13.n1.2020.29582.
- BORKO, H. Information Science: what is it? *American Documentation*, 1968, vol. 19, nº 1, p. 3-5. Disponível em: doi: 10.1002/asi.5090190103.
- CURTY, R.G. y DELBIANCO, N.R. As diferentes metrias dos estudos métricos da informação: evolução epistemológica, inter-relações e representações. *Enc.Bibli. R. Eletr.Bibliotecon.Ci.Inf.*, 2020, vol. 25, p. 1-21. Disponível em: doi: 10.5007/1518-2924.2020.e74593.
- DIMENSIONS. *Why did we build Dimensions?* Disponível em: <<https://www.dimensions.ai/why-dimensions/>> [Acesso em: 3 jan. 2021]
- FONSECA, E.N. de. Bibliografia estatística e bibliometria: uma reinvidicação de prioridades. *Ciência da Informação*, 1973, vol. 2, nº 1, p. 5-7. Disponível em: <<http://revista.ibict.br/ciinf/article/view/19/19>> [Acesso em: 4 jan. 2021]

- GLÄNZEL, W. *Bibliometrics as a research field: a course on theory and application of bibliometric indicators*. Bélgica, 2003. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/download?doi=10.1.1.97.5311&rep=rep1&type=pdf>> [Acesso em: 28 set. 2021]
- GRÁCIO, M.C.C. *Análises relacionais de citação para a identificação de domínios científicos*. São Paulo: Cultura Acadêmica, 2020. Disponível em: <https://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/book/166> [Acesso em: 10 jul. 2020]
- HARZING, A.W. Two new kids on the block: how do crossref and Dimensions compare with Google Scholar, Microsoft Academic, Scopus and the Web of Science? *Scientometrics*, 2019, vol. 120, nº 1, p. 341-349. Disponível em: doi: 10.1007/s11192-019-03114-y.
- HARZING, A.W. y ALAKANGAS, S. Google Scholar, Scopus and the Web of Science: a longitudinal and cross-disciplinary comparison. *Scientometrics*, 2016, vol. 106, nº 2, p. 787–804. Disponível em: doi: 10.1007/s11192-015-1798-9.
- HJØRLAND, B. Domain analysis in information science: eleven approaches: traditional as well as innovative. *Journal of Documentation*, 2002, vol. 58, nº 4, p. 422–462. Disponível em: doi: 10.1108/00220410210431136.
- KOLTAY, T. Library and information Science and the digital humanities. *Journal of Documentation*, 2016, vol. 72, nº 4, p. 781-792. Disponível em: doi: 10.1108/JD-01-2016-0008.
- MACIAS-CHAPULA, C.A. O papel da informetria e da cienciométrica e sua perspectiva nacional e internacional. *Ciência da Informação*, 1998, vol. 27, nº 2, p. 134-140. Disponível em: doi: 10.1590/S0100-19651998000200005.
- MOED, H.F. *Applied evaluative informetrics*. Cham, Switzerland: Springer, 2017. Disponível em: doi: 10.1007/978-3-319-60522-7_14.
- OLIVEIRA, E.F.T. de. *Estudos métricos da informação no Brasil: indicadores de produção, colaboração, impacto e visibilidade*. Marília, SP: Oficina Universitária, 2018. Disponível em: <http://ebooks.marilia.unesp.br/index.php/lab_editorial/catalog/book/2> [Acesso em: 10 jun. 2020]
- OLIVEIRA, E.F.T. y GRÁCIO, M.C.C. *Questões métricas em organização da informação*. Marília, SP: PPGCI/UNESP, 2016.
- OLIVEIRA, M. de. *Ciência da informação e biblioteconomia: novos conteúdos e espaços de atuação*. Belo Horizonte: Ed. UFMG, 2005.
- ORDUÑA-MALEA, E. y DELGADO-LÓPEZ-CÓZAR, E. Dimensions: re-discovering the ecosystem of scientific information. *El Profesional de la Información*, 2018, vol. 27, nº 2, p. 420-431. Disponível em: doi: 10.3145/epi.2018.mar.21.
- PIMENTA, R.M. Por que humanidades digitais na ciência da informação? perspectivas progressas e futuras de uma prática transdisciplinar comum. *Informação & Sociedade: Estudos*, 2020, vol. 30, nº 2, p. 1-20. Disponível em: doi: 10.22478/ufpb.1809-4783.2020v30n2.52122.
- PIMENTA, R.M. Os objetos técnicos e seus papéis no horizonte das humanidades digitais: um caso para a ciência da informação. *Revista Conhecimento em Ação*, 2016, vol. 1, nº 2. Disponível em: doi: 10.47681/rca.v1i2.5525.
- ROBINSON, L.; PRIEGO, E. y BAWDEN, D. Library and Information Science and Digital Humanities: Two Disciplines, Joint Future? In: PEHAR, F.; SCHLÖGL, C.; WOLFF, C. (Eds). *Re: inventing Information Science in the Networked Society, Proceedings of the 14th International Symposium on Information science (ISI2015)*, Zadar, Croatia, 19th-21st May 2015. Verlag Werner Hülsbusch, Glückstadt, p. 44-54. Disponível em: <http://doi.org/10.5281/zenodo.17969>.
- RUSSELL, I.G. ¿Qué son las Humanidades Digitales? *Revista Digital Universitária*, 2011, vol. 12, nº 7, p. 3-8. Disponível em: <<http://www.revista.unam.mx/vol.12/num7/art68/index.html>> [Acesso em: 21 abr. 2021]
- SARACEVIC, T. Ciência da Informação: origem, evolução e relações. *Perspectiva em Ciência da Informação*, 1996, vol. 1, nº 1, p. 41-62. Disponível em: <<http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/235>> [Acesso em: 20 jan. 2021]
- SCIMAGO. SJR – SCImago Journal & Country Rank. Disponível em: <<http://www.scimagojr.com>> [Acesso em: 30 mar. 2021]
- SIQUEIRA, M.N. de y FLORES, D. Ciência da informação e humanidades digitais: diálogos possíveis de uma relação em desenvolvimento – artigos científicos no Brasil. *Liinc em Revista*, 2019, vol. 15, nº 1, p. 85-107. Disponível em: doi: 10.18617/liinc.v15i1.4563.
- SUGIMOTO, C.R. y LARIVIÈRE, V. *Measuring research: what everyone needs to know*. Oxford: Oxford University Press, 2018.
- SULA, C.A. Digital Humanities and Libraries: a conceptual model. *Journal of Library Administration*, 2013, vol. 53, nº 1, p. 10-26. Disponível em: <http://doi.org/10.1080/01930826.2013.756680>.
- VANTI, N.A.P. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. *Ciência da Informação*, 2002, vol. 31, nº 2, p. 152-162. Disponível em: doi: 10.1590/S0100-19652002000200016.