

**PRODUÇÃO CIENTÍFICA SOBRE HUMANIDADES DIGITAIS E CIÊNCIA ABERTA
NA WEB DA CIÊNCIA (WOS) E BANCOS DE DADOS SCOPUS 2014-2020**

**LA PRODUCCIÓN CIENTÍFICA SOBRE HUMANIDADES DIGITALES Y CIENCIA
ABIERTA EN LAS BASES DE DATOS WEB OF SCIENCE (WOS) Y SCOPUS 2014-
2020**

**THE SCIENTIFIC PRODUCTION ON DIGITAL HUMANITIES AND OPEN SCIENCE
ON THE WEB OF SCIENCE (WOS) AND SCOPUS DATABASES 2014-2020**

Fabiane Führ

Universidade Federal de Santa Catarina (UFSC), Brasil

fabifuhr@gmail.com

<https://orcid.org/0000-0002-3723-050X>

Edgar Bisset Alvarez

Universidade Federal de Santa Catarina (UFSC), Brasil

edgar.bisset@ufsc.br

<https://orcid.org/0000-0002-5388-5944>

Recibido: 16 de julio de 2021

Revisado: 30 de julio de 2021

Aprobado: 8 de septiembre de 2021

Cómo citar: Führ, F; Bisset Alvarez, E. (2021). La producción científica sobre humanidades digitales y ciencia abierta en las bases de datos Web of Science (WoS) y Scopus (2014-2020). *Bibliotecas. Anales de Investigación*; 17(3), 1-24

RESUMO

Objetivo. Identifique quais aspectos da ciência aberta são abordados nas publicações com foco nas humanidades digitais. **Desenho / Metodologia / Abordagem.** Identifica a produção científica em humanidades digitais e ciências abertas nas bases de dados Web of Science (WoS) e Scopus. Ele usa o gerenciador bibliográfico Zotero para organizar os dados bibliográficos e aplicar os resultados às

ferramentas de mineração de dados Sobek e Voyant Tools. **Resultados / discussão.** Dos 17 artigos analisados, 6 foram escritos em inglês e 5 em alemão. Os anos mais produtivos foram 2019 e 2020 com 5 publicações. A revista com mais publicações sobre o assunto é *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare (VÖB)* da Áustria. Apenas 7 periódicos são classificados pelo Qualis Capes, dos quais 5 estão no estrato A. Apenas 5 artigos não utilizam projetos ou programas voltados para as humanidades digitais para suas discussões. Todos os artigos abordam Humanidades Digitais e ações para a Ciência Aberta, mas alguns enfatizam a Ciência Aberta e os desafios que ela impõe. **Conclusões.** Mostra a importância do gerenciamento de dados e a necessidade de documentos orientadores. Também aponta a relevância do padrão de metadados, para tornar os dados mais adequados aos princípios do FAIR, além de formar pesquisadores e cidadãos para promover a colaboração entre instituições e pessoas com diversas formações para a valorização da ciência aberta. **Originalidade / valor.** Percepção de pesquisadores sobre a temática da ciência aberta nas Humanidades Digitais.

PALAVRAS-CHAVE: Humanidades Digitais; Ciência aberta; Mineração de dados.

RESUMEN

Objetivo. Identificar qué aspectos de la ciencia abierta se abordan en las publicaciones dirigidas a las humanidades digitales. **Diseño / Metodología / Enfoque.** Identifica la producción científica sobre humanidades digitales y ciencia abierta en las bases de datos Web of Science (WoS) y Scopus. Utiliza el administrador bibliográfico de Zotero para organizar los datos bibliográficos y aplicar los resultados a las herramientas de minería de datos Sobek y Voyant Tools. **Resultados / Discusión.** De los 17 artículos analizados, 6 fueron escritos en inglés y 5 en alemán. Los años más productivos fueron 2019 y 2020 con 5 publicaciones. La revista con más publicaciones sobre el tema es *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare (VÖB)* de Austria. Solo 7 revistas están clasificadas por Qualis Capes, de las cuales 5 están en el estrato A. Solo 5 artículos no utilizan proyectos o programas dirigidos a las humanidades digitales para sus discusiones. Todos los artículos abordan las Humanidades Digitales y las acciones para la Ciencia Abierta, pero algunos enfatizan la Ciencia Abierta y los desafíos que impone. **Conclusiones.** Muestra la importancia de la gestión de datos y la necesidad de documentos guía. También señala la relevancia del estándar de metadatos, para adecuar los datos a los principios FAIR, además de capacitar a investigadores y ciudadanos para promover la colaboración entre instituciones y personas de diversos orígenes para valorar la ciencia abierta. **Originalidad / Valor.** Percepción de los investigadores sobre el tema de la ciencia abierta en Humanidades Digitales.

PALABRAS CLAVE: Humanidades digitales; Ciencia Abierta; Procesamiento de dato.

ABSTRACT

Objective. Identify which aspects of open science are covered in publications focused on the digital humanities. **Design/Methodology/Approach.** It identifies scientific production on digital humanities and open science in the Web of Science (WoS) and Scopus databases. It uses the Zotero bibliographic manager to organize bibliographic data and apply the results to the data mining tools Sobek and Voyant Tools. **Results/Discussion.** Of the 17 articles analyzed, 6 were written in English and 5 in German. The most productive years were 2019 and 2020 with 5 publications. The journal with the most publications on the subject is *Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare (VÖB)*

from Austria. Only 7 journals are classified by Qualis Capes, of which 5 are in stratum A. Only 5 articles do not use projects or programs aimed at the digital humanities for their discussions. All articles address Digital Humanities and actions for Open Science, but some emphasize Open Science and the challenges it imposes. **Conclusions.** Shows the importance of data management and the need for guiding documents. It also points out the relevance of the metadata standard, to make data more adequate to FAIR principles, in addition to training researchers and citizens to promote collaboration between institutions and people with diverse backgrounds to value open science. **Originality/Value.** Perception of researchers on the theme of open science in Digital Humanities.

KEYWORDS: Digital Humanities; Open Science; Data mining.

INTRODUÇÃO

O termo *Digital Humanities* ganhou visibilidade com a publicação do livro *A companion to Digital Humanities*, editado por Susan Schreibman, Ray Siemens e John Unsworth, lançado em 2004. Todavia, sua origem é anterior e remonta a pesquisas desenvolvidas nas décadas de 1940 e 1950, pelo Padre Roberto A. Busa que criou o *Index Thomisticus*, que é considerada a primeira aplicação da computação aos estudos linguísticos nas obras de São Tomás de Aquino (Guerreiro, 2020).

Segundo Dalbello (2011) a linha do tempo das Humanidades Digitais começa com a introdução de métodos computacionais nos campos literário, filológico e filosófico na década de 1950 e continua com foco nos arquivos multimídia pesquisáveis e nos textos estruturados nas décadas de 1980 e 1990, além da edição de textos de projetos e coleções eletrônicas das décadas passadas [anos 2000].

O campo de pesquisas em Humanidades Digitais tem aumentado exponencialmente devido a crescente digitalização de documentos, que tem gerado uma grande massa de dados, ampliado o número de pesquisas colaborativas e alterado a cadeia de produção científica. Contudo, os projetos de Humanidades Digitais não se restringem apenas a acessibilidade e disseminação do conhecimento, mas dizem respeito também às formas de criação e de divulgação dos mesmos (Almeida & Damian, 2017).

As Humanidades Digitais permitem a criação interdisciplinar, intercultural e colaborativa de conhecimentos, bem como a sua acessibilidade e socialização, permitindo que os resultados de pesquisa cheguem a mais pessoas. Os métodos digitais têm proporcionado espaços de construção e participação cidadã que tornam possível desenvolver e gerar soluções para questões complexas de pesquisas (Londoño, 2020).

Essa mesma cadeia de produção científica tem sofrido modificações devido ao movimento em favor da Ciência Aberta. Este movimento tem se consolidado em muitos países a partir do movimento em prol do Acesso Aberto, cujo marco decisivo foi a Declaração de Budapeste publicada em 2002. A Ciência Aberta tem como propósito o compartilhamento e acesso a publicações e dados de pesquisas, promove a abertura do processo científico, auxilia na transferência do conhecimento, amplia os impactos sociais e econômicos da ciência e reforça a responsabilidade social da ciência (Santos, Almeida, Henning, 2017).

Assim, a Ciência Aberta é um movimento que propõe novas formas de produção colaborativa, interativa e compartilhada da informação, do conhecimento e da cultura (Albagli, 2017). Esse movimento tem ganhado cada vez mais visibilidade e se estendido às mais diversas áreas de investigação.

Diante desse contexto, é importante observar nas publicações sobre Humanidades Digitais como a temática da ciência aberta tem se desenvolvido. Portanto, o objetivo geral deste artigo é identificar quais aspectos da Ciência Aberta estão sendo abordados nas publicações referentes as Humanidades Digitais, nas bases de dados Web of Science (WoS) e Scopus.

Os objetivos específicos são: a) identificar a produção científica sobre Ciência Aberta e Humanidades Digitais indexada nas bases de dados Web of Science e Scopus de 2014 a 2020; b) aplicar os textos do corpus nos mineradores de textos: *Sobek*, da Universidade Federal do Rio Grande do Sul e no *Voyant Tools*, aplicativo desenvolvido por professores canadenses e c) descrever como a Ciência Aberta é abordada em cada um dos artigos do corpus da pesquisa e de que maneira estes se relacionam com as Humanidades Digitais.

Acredita-se que um dos motivadores para o crescimento de pesquisas e publicações sobre a Ciência Aberta tenha sido o desenvolvimento do movimento em prol do acesso aberto ao conhecimento, que teve na Declaração de Budapeste, de 2002, seu marco inicial. Todavia, todas as mudanças requerem um período para seu amadurecimento, o que pode ser observado nesta pesquisa que se propõe a analisar quando e como a ciência aberta passou a figurar entre as propostas e projetos relacionados as Humanidades Digitais.

METODOLOGIA

Com o objetivo de identificar a produção científica na área de Humanidades Digitais e relacioná-la com a Ciência Aberta, definiu-se como estratégia de busca a expressão [("humanidades digitais" OR "humanidades digitales" OR "digital humanities") AND ("ciência aberta" OR "ciencia abierta" OR "open science")].

A estratégia foi aplicada nas bases de dados Scopus, Web of Science (WoS) no dia 18 de agosto de 2021. A busca foi realizada nos campos título, palavra-chave e resumo. Não houve aplicação de filtro com relação a data de publicação, pois trata-se de um tema recente e dessa forma, seria possível obter um número maior de registros para análise. Todavia, sobre esses resultados foi aplicado filtro para que fossem recuperados apenas os artigos científicos. Dessa forma, foram recuperados XX artigos, conforme apresentado no Quadro 1.

Quadro 1. Pesquisa realizada em bases de dados

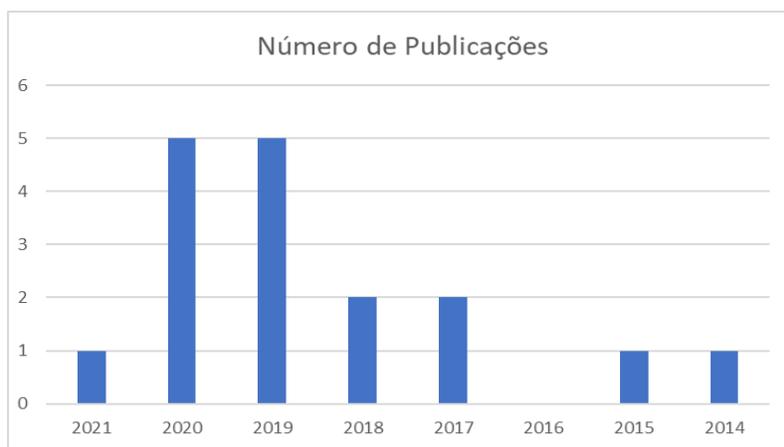
Scopus	12 artigos	2014 a 2021
Web of Science	18 artigos	2014 a 2021
Total	30 artigos	2014 a 2021

O gerenciador bibliográfico Zotero foi utilizado para a coleta, o armazenamento e a organização dos artigos recuperados no levantamento bibliográfico. Após a organização, foram excluídos 9 arquivos duplicados e 4 artigos que não tinham relação com o escopo da pesquisa. Restando assim, 17 artigos para análise.

Apresentação e discussão dos resultados

A pesquisa foi realizada em 17 artigos, dos quais 6 foram escritos em língua inglesa e outros 5, em língua alemã. 3 artigos foram redigidos em português e 2 em espanhol. Um dos artigos publicados era bilíngue, inglês/português. Com relação a data de publicação, 2019 e 2020 foram os anos com maior número de publicações. 2021 apresenta apenas 1 publicação, porém as revistas continuam recebendo submissões e pode ser que haja artigos sobre a temática sendo editados.

Gráfico 1. Número de publicações por ano



Já em relação ao país de publicação do periódico observou-se que a Áustria e o Brasil são os países mais produtivos com 3 publicações, como pode ser observado no gráfico 2.

Gráfico 2. Países por publicação



Os periódicos foram analisados com relação ao Qualis Capes (Quadriênio 2013-2016). Com relação ao fator de impacto (Journal Impact Factor) fornecido pelo Journal Citation Reports, da Web of Science Group e os periódicos que não foram identificados no Journal Impact Factor, foram analisados para verificar se estavam na lista do Emerging Source Citation Report, da Web of Science Group, ambos Clarivate Analytics (Web of Science Group, 2020). Com relação ao CiteScore, que é uma métrica que fornece a indicação de impacto da pesquisa. O CiteScore é desenvolvido pela Scopus, base de dados da Elsevier (Scopus, 2019).

Tabela 1. Análise dos periódicos

ASLIB Journal of Information Management	1	*	1.903	*	3.3
Bibliothek - Forschung und Praxis	1	*	*	0.10	*
Cataloging and Classification Quarterly	1	A2	*	0.22	1.6/2015
Digital Humanities Quarterly	1	*	*	1.23	0.2/2019
ERCIM News	1	*	*	0.04	*
Estudos Históricos	2	B1	*	0.81	0.2
História, Ciências e	1	A2	*	0.26	

¹ O Qualis do periódico Literary and Linguistic Computing refere-se a área de avaliação Linguística e Literatura. Os demais pertencem a área Comunicação e Informação

Saúde - Manguinhos					
Information Research	1	A1	0.780		1.9
Literary and Linguistic Computing	1	A1	1.125/2016		*
Mitteilungen der Vereinigung Österreichischer Bibliothekarinnen & Bibliothekare (VÖB)	3	*	*	*	*
Profesional de la Informacion	1	A1	2.253	*	3.1
Publications	1	B5	*	0.79	2.6
Revista General de Información y Documentación	1	*		0.11	0.6
Zeitschrift für Germanistik	1	*	*	1.26	0.2/2015

* Não havia informações sobre os periódicos nas fontes consultadas.

Observa-se que dos 17 periódicos, 5 encontram-se no estrato A, 2 no estrato B e 7 não se encontram em nenhum estrato. Observa-se que nenhum periódico em língua alemã apresenta informações no Qualis Capes, o que pode ser justificado pela barreira linguística, uma vez que o Qualis Capes reflete onde os docentes e pesquisadores da área tem publicado os resultados de suas pesquisas (Capes, 2021).

Todos os artigos foram analisados quanto a sua inserção na área de Humanidades Digitais e na Ciência Aberta. Assim, a tabela 2 apresenta os 17 artigos selecionados, seus autores, bem como os projetos e programas relativos as Humanidades Digitais que são descritos e abordados na publicação.

Tabela 2. Autoria e projetos relativos as Humanidades Digitais

2021	Cantón, CIM	Un análisis cuantitativo del acceso abierto en las revistas de Humanidades Digitales	
2020	Ferla, L; Lesser, J; Ferreira, KR; Miyasaka, C; Atique, F; Musa, D; Britt, AG; Rogers, TD; Fook, KD;	Pauliceia 2.0: mapeamento colaborativo da história de São Paulo, 1870-1940	Pauliceia 2.0

	Vijaykumar, N.		
2020	Hocker, J; Schindler, C; Rittberger, M.	Participatory design for ontologies: a case study of an open science ontology for qualitative coding schemas	
2020	Albagli, S; Py, H; Iwana, AY	Geovisualização de dados e a ciência aberta e cidadã – a experiência da Plataforma LindaGeo	Projeto de pesquisa-ação Ciência Aberta Ubatuba
2020	Rollo, MF	Desafios e responsabilidades das humanidades digitais: preservar a memória, valorizar o patrimônio, promover e disseminar o conhecimento. O programa Memória para Todos	Programa Memória de Todos
2020	Ferla, LAC; Lima, LFS; Feitler, B	Novidades no front: experiências com humanidades digitais em um curso de história na periferia da Grande São Paulo	Projeto Implementação da tecnologia de Sistemas de Informações geográficas (SIG) em investigações históricas (2012-2013); Projeto Pauliceia 2.0; Projeto de avaliação e edição de verbetes da Enciclopédia Virtual Wikipédia
2019	Lahti, L; Marjanen, J; Roivainen, H; Tolonen, M	Bibliographic Data Science and the History of the Book (c. 1500–1800)	Bibliographic data science
2019	Knöchelmann, M	Open Science in the Humanities, or: Open Humanities?	
2019	Anglada, LM	Muchos cambios y algunas certezas para las bibliotecas de investigación, especializadas y centros de documentación	Coleção do Museu do Prado; Transcribe Betham; eBird e Mapa Literali Catalã
2019	Stigler, J; Klug, HW	Konde – ein netzwerk bringt forschungs- und glam-institutionen zusammen. Ein projektbericht	Projeto Konde (acrônimo de Kompetenznetzwerk Digitale Edition)
2019	Wutke, U; Spiecker, C; Neuroth, H	PARTHENOS – Eine digitale Forschungsinfrastruktur für die Geistes- und Kulturwissenschaften	Projeto PARTHENOS (acrônimo de Pooling Activities, Resources and Tools for Heritage E-research Networking, Optimization and Synergies)
2018	Hagmann, D	Überlegungen zur nutzung von phaidra als repositorium für digitale archäologische daten	Repositório PHAIDRA e escavação "Molino San Vincenzo"

2018	Steinerova, J	Perceptions of the information environment by researchers: a qualitative study	
2017	Bassett, S; Di Giorgio, S; Ronzino, P	A Data Management Plan for Digital Humanities: the PARTHENOS Model	Projeto PARTHENOS
2017	Baum, C	Digital gap oder digital turn? Literaturwissenschaft und das digitale Zeitalter	
2015	Blumesberger, S	Die welt der Metadaten im universum von repositorien	Projekts e-Infrastructures Austria
2014	Wells, JJ; Kansa, EC; Kansa, SW; Yerka, SJ; Anderson, DG; Bissett, TG; Myers, KN; Demuth, RC	Web-based discovery and integration of archaeological historic properties inventory data: The Digital Index of North American Archaeology (DINAA)	The Digital Index of North American Archaeology (DINAA)

Com relação ao tipo de autoria, observou-se que 08 artigos são de autoria única. Os demais artigos possuem autoria compartilhada distribuída da seguinte forma: 5 artigos possuem 3 autores e apenas um artigo publicado por 2 autores, 4 autores, 8 autores e 10 autores, respectivamente.

Com relação aos projetos e programas, apenas 05 artigos (Cantón, 2021; Hocker, Schindler & Rittberger, 2020; Knöchelmann, 2019; Steinerova, 2018; Baum, 2017) não utilizam projetos ou programas em Humanidades Digitais para suas discussões. Todavia, apresentam questões relevantes quanto a temáticas das humanidades digitais e da ciência aberta.

A mineração nos textos

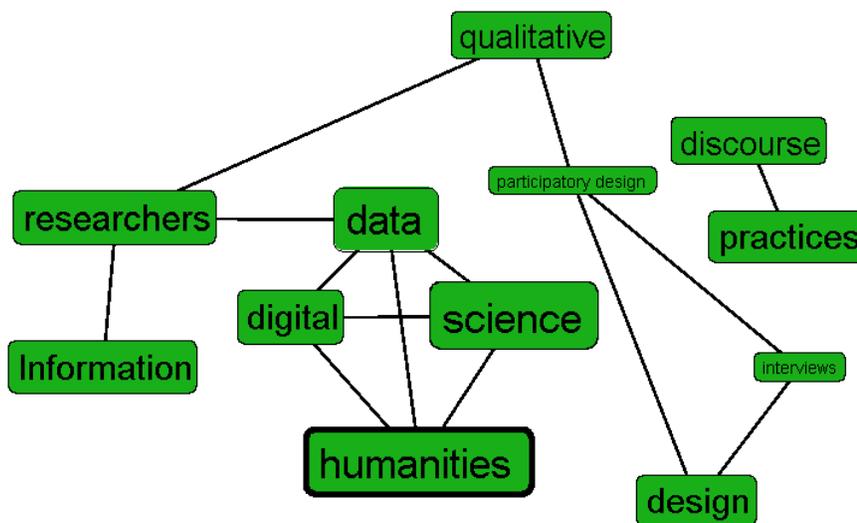
Com o intuito de identificar informações relevantes nos textos fez-se uma análise dos resumos utilizando os mineradores de texto *Sobek* e *Voyant Tools*. Os resumos dos textos foram reunidos por idioma. Dessa forma, obteve-se 4 conjuntos de textos em arquivos .txt: Alemão (reuniu em um único documento os 5 resumos em alemão); Espanhol (reuniu 2 resumos em espanhol); Inglês (reuniu os 6 resumos em inglês) e Português (reuniu os 4 textos em português, uma vez que se optou por incluir o resumo do texto bilíngue com os resumos de língua portuguesa).

A ferramenta *Sobek* foi desenvolvida a partir de um algoritmo inicialmente criado por Schenker, em 2003, e modificado pelo Grupo de Pesquisa GTech.Edu que o adequou às aplicações educacionais (Gtech.Edu, 2020). A ferramenta está adaptada apenas aos idiomas inglês e português, assim ao aplicar os textos em alemão e em espanhol os resultados foram inconsistentes, uma vez que as palavras mais frequentes consideradas pelo sistema foram conectivos.

Por isso, buscou-se outra ferramenta de mineração de texto que fosse de acesso aberto e que permitisse a mineração textual em outros idiomas. A ferramenta escolhida foi o *Voyant Tools*, que é um aplicativo online, de código aberto que permite análise de textos e que foi desenvolvido pelos professores canadenses Stéfán Sinclair e Geoffrey Rockwell (Larhud, 2020).

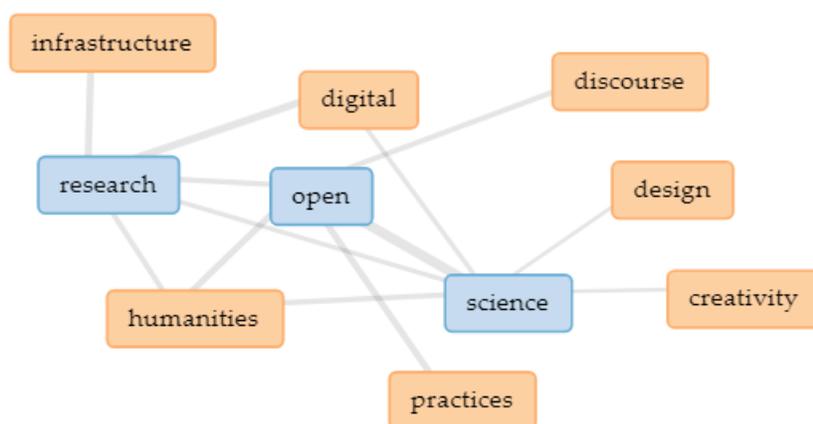
Ao aplicar os textos em Inglês no software *Sobek*, obtive o seguinte resultado:

Figura 1. Resultado Mineração *Sobek* Inglês



É possível observar que as palavras de maior destaque são *data*, *humanities*, *science* e *digital*. Essas palavras apresentam ligações entre si, mostrando o quão ligadas estão no contexto dessa pesquisa. Outros termos em evidência são *researchers* (pesquisadores) que está diretamente ligado ao conjunto anterior, bem como aos termos *information* (informação) e *qualitative* (qualitativa), que pode evidenciar a busca dos pesquisadores, bem como as pesquisas realizadas no âmbito desta pesquisa.

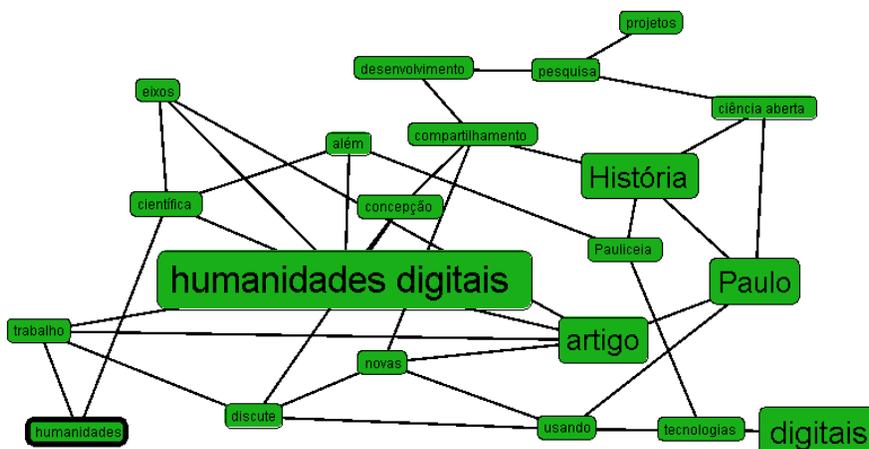
Figura 2. Resultado mineração *Voyant Tools* Inglês



Ao aplicar o mesmo texto no *Voyant Tools* observa-se que as palavras-chave são representadas em azul (Figura 2). As palavras mais relevantes são *research* (pesquisa, como 24 ocorrências), *open* (aberto, com 22 ocorrências) e *science* (ciência, como 16 ocorrências). A ligação mais forte (representada pela linha mais espessa) ocorre entre *open* e *science*, justamente termos da pesquisa. Observa-se que o termo *research* está vinculado a *open*, *science*, *infraestrutura*, *digital*, *humanities* enfatizando as infraestruturas para pesquisas em humanidades digitais. O termo *open* está ligado a *research*, *science*, *humanities*, *discourse* e *practices*, evidenciando a busca pelos discursos e práticas abertas nas humanidades digitais. E o termo *science* está vinculado as palavras *research*, *open*, *digital*, *design* e *creativity* ressaltando a concepção dos projetos criativos no âmbito das humanidades digitais e da ciência aberta.

Observou-se que no software *Voyant Tools* a ligação entre os termos *open* e *science*, que é um dos objetos desta pesquisa fica evidente, porém a relação entre *digital* e *humanities* já não é tão aparente. Enquanto isso, na análise produzida pelo software *Sobek*, é possível observar a relação entre *digital*, *humanities* e *science*, porém o termo *open* nem figura entre os termos relacionados pelo sistema.

Figura 3. Resultado mineração *Sobek* Português



Na figura 3, que representa a mineração do texto em língua portuguesa no software *Sobek*, observa-se a prevalência do termo *humanidades digitais* e sua ligação com os termos *artigo*, *trabalho*, *científica*, *eixos*, *além* e *concepção*. O termo *história* também apresenta grande evidência e relação com a expressão *Paulo*, que se refere a cidade de São Paulo, tema de uma das pesquisas analisadas. Outro termo em evidência é *digitais* que está ligado a *tecnologias* e a *Pauliceia*, novamente relacionando-o com a pesquisa sobre a história da cidade de São Paulo. A expressão *ciência aberta* está relacionada ao termo *história* e a *pesquisa*, que por sua vez está relacionado com *projetos* e *desenvolvimento*.

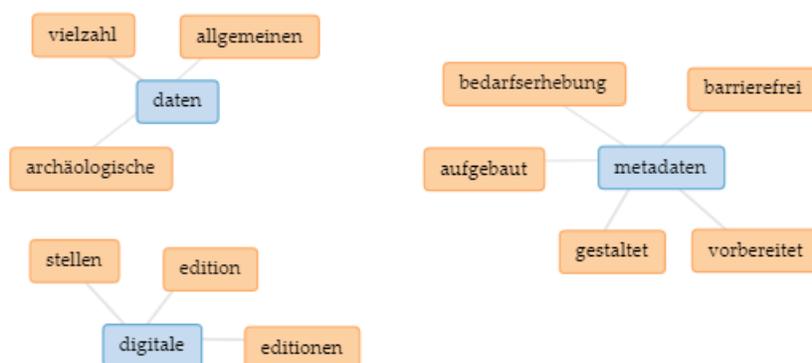
Figura 4. Resultado mineração *Voyant Tools* Português – Nuvem de palavras



Outra forma utilizada para analisar o resultado foi por meio da nuvem de palavras no *Voyant Tools* (Figura 4) apresenta os termos *digitais* e *humanidades* em destaque. Estes termos apareceram 11 e 9 vezes, respectivamente, nos resumos reunidos e analisados em português. Outras palavras como *artigo*, *história*, *Paulo* e *projeto* tiveram um destaque de 3 incidências no resumo analisado. Outras palavras podem ser visualizadas e mostram a diversidade que podemos obter em projetos de dessa natureza, discutindo o compartilhamento, a circulação, os dados, os projetos, a pesquisa, a ciências, o fazer colaborativo, as tecnologias, entre outros.

A figura 5 traz como resultado a mineração nos resumos em língua alemã. O que se observa é que as palavras-chave representadas em azul, não apresentam uma ligação entre si. A palavra-chave é *daten* que significa dados coocorre com as *vielfalt* (grande variedade, multiplicidade) *allgemeinen* (geral, universal) e *archäologische* (Arqueologia, arqueológica). A palavra-chave *digitale* (digital, digitais) está associada a *stellen* (perguntar, colocar), *edition* (edição) e *editionen* (edições). A palavra-chave *metadaten* (metadados) coocorre com as palavras *bedarfshebung* (avaliação/identificação das necessidades), *barrierefrei* (acessíveis, acessibilidade), *aufgebaut* (construído, criado, desenvolvido), *gestaltet* (estruturado) e *vorberitet* (preparado, pronto). O que se observa é que a partir do grafo das palavras-chave e suas coocorrências não é possível estabelecer uma relação entre as Humanidades Digitais e a Ciência Aberta.

Figura 5. Resultado mineração *Voyant Tools* Alemão



Optou-se assim por extrair a nuvem de palavras utilizando o software *Voyant Tools*. A partir dessa imagem (Figura 6) observa-se que as palavras *digitale* e *metadaten* são as que aparecem com mais evidência e palavras importantes para a temática desta pesquisa como *daten*, *open* e *science* aparecem com mais clareza. Todavia, as Humanidades Digitais não estão expressas na nuvem de palavras.

Figura 6. Resultado mineração *Voyant Tools* Alemão – Nuvem de palavras



Para o resumo em espanhol optou-se apenas pela nuvem de palavras que apresenta os termos mais relevantes no texto, tais como: *información* (7 ocorrências), *acceso* (6), *centros* (5), *digitales* (5), *humanidades* (5), *publicación* (5), *revistas* (4), *abierto* (3), *cambios* (3), *corpus* (3), *estudio* (3), *publicaciones* (3), *abierta* (2), *bibliotecas* (2), entre outras com menor relevância. Nesta nuvem de palavras (Figura 7) observa-se as Humanidades Digitais em evidência, já a Ciência Aberta não é tão aparente mas está presente diferente de outras nuvens anteriormente observadas.

Fig. 7. Resultado mineração *Voyant Tools* Espanhol – Nuvem de palavras

interdisciplinar que permita que os pesquisadores de humanidades tenham acesso aos dados, as ferramentas e aos serviços baseados em políticas, diretrizes e padrões comuns (Basset, Di Giorgio & Ronzino, 2017).

Em 2019, Wuttke, Spiecker e Neuroth também trouxeram ações sobre o Projeto PARTHENOS. Segundo os autores as áreas de maior importância para o projeto são definir como se dará o arquivamento a longo prazo, a gestão da propriedade intelectual, sua implementação, o desenvolvimento de padrões, serviços e métodos comuns para o uso interdisciplinar e subsequente uso de dados, bem como treinamento e educação adicional. Ao longo do artigo são apresentadas as ações desenvolvidas em pacotes de trabalho, sendo que os pacotes mais relevantes são os pacotes 4, 7 e 8. O pacote 4 desenvolveu um kit de padronização, que consiste em uma documentação sobre os padrões relevantes para as ciências humanas e culturais em meio digital. Este pacote também pode ser utilizado para o ensino, e em parceria com os pacotes 7 e 8 desenvolve materiais de comunicação e treinamento sobre a importância da padronização para o processo de pesquisa (Wuttke, Spiecker & Neuroth, 2019).

No ano de 2018, Hagmann apresenta o repositório PHAIDRA, da Universidade de Viena, que armazena a longo prazo dados digitais de pesquisas científicas em um ambiente com padrões sustentáveis. Por isso, o autor considera o PHAIDRA ideal para o armazenamento de dados arqueológicos, como os dados que tem sido coletados no Sítio Molino San Vincenzo na Toscana, Itália. A Universidade de Viena tem realizado investigações arqueológicas no sítio, desde 2012. Há dados coletados em formato digital e analógico. Estes, são digitalizados para posterior processamento. O autor enfatiza o uso de tecnologias de informação e comunicação (TIC) na arqueologia, o que é um pré-requisito para a geração de dados digitais em arqueologia (Hagmann, 2018).

Retomando ao ano de 2019, Lathi et al. apresentam o projeto *Bibliographic data science*, no qual analisam a ciência de dados bibliográficos como estudo que deriva da área de ciência de dados. Para a realização da análise utilizam as bibliografias nacionais suecas (FBN e SNB), o catálogo inglês de títulos curtos (ESTC) e o banco de dados do patrimônio do livro impresso (HPBD). Essas bibliografias cobrem mais de 6 milhões de entradas de produtos impressos na Europa e em outros lugares. Segundo os autores com este estudo é possível demonstrar, por meio de uma abordagem qualitativa, a história do livro (Lathi et al., 2019).

Segundo os autores os esforços para a harmonização automatizada em grande escala podem melhorar a confiabilidade geral e a comensurabilidade entre as coleções de metadados, complementando com o LOD e outras tecnologias que se concentram no gerenciamento e distribuição de dados. A ciência de dados bibliográficos tem como objetivo preencher uma lacuna importante da área, visto que muitos metadados bibliográficos apresentam entradas imprecisas, vieses de coleta e informações ausentes (Lathi et al., 2019).

Para Anglada (2019) as Humanidades Digitais são um movimento que afeta a vida e a evolução dos centros de informação. Segundo o autor a sociedade se aproxima do momento onde aquilo que não está na rede já não tem importância, por isso é fundamental que as bibliotecas e centros de documentação desloquem sua atenção para o documento de forma que este tenha seu acesso facilitado. Além disso, o

autor sugere que se fomente a participação dos cidadãos, como ocorre nas iniciativas: Coleção do Museu do Prado, o projeto *Transcribe Betham*, o projeto *eBird* e o Mapa *Literali Catalã* (Anglada, 2019).

O autor sugere que é possível observar claramente os efeitos da ciência aberta sobre as Humanidades Digitais, uma vez que pesquisadores necessitam que os objetos digitais estejam acompanhados de informações sobre os processos que os tornaram legíveis pelo computador, para que possam ser compartilhados e reutilizados por diferentes grupos posteriormente (Anglada, 2019).

Stiegler e Klug (2019) expõem o Projeto Konde (acrônimo de *Kompetenznetzwerk Digitale Edition*) que define os pré-requisitos para o estabelecimento de uma plataforma para edições digitais, que tem como objetivo o desenvolvimento e a preservação do patrimônio cultural. Os autores afirmam que as edições digitais são um produto da disciplina digital de Humanidades, em que métodos auxiliados por computador são usados para criar, pesquisar e disseminar publicações de fontes científicas confiáveis (Stiegler & Klug, 2019).

Com o projeto Konde, a Austrália pretende tornar-se líder em inovação em edições digitais. Para isso conta com a participação de pesquisadores e profissionais de bibliotecas, arquivos e museus para fortalecer os centros de pesquisas e efetivar a colaboração entre os participantes (Stiegler & Klug, 2019).

Memória de Todos é o programa apresentado por Rollo (2020) e tem como objetivo promover a educação patrimonial, a literacia, o desenvolvimento de competências digitais e a democratização do acesso às ferramentas de investigação histórica, além de fomentar a recolha, a partilha e a preservação de memórias e de testemunhos.

Para a autora as Humanidades Digitais compreendem desde o uso de ferramentas, a pesquisa em informação, a organização, o depósito de conteúdos, ou inclusive a programação e utilização de bases de dados, ou ferramentas de computação em sua totalidade. Todavia, essas competências são escassas na própria comunidade de humanidades. Ainda segundo a autora, o digital trouxe uma série de oportunidades, no sentido da ampliação dos conteúdos, em especial os conteúdos imateriais que podem ser registrados e preservados para as futuras gerações (Rollo, 2020).

Os autores Ferla, Lima e Feitler (2020) apresentam o Projeto Implementação da tecnologia de Sistemas de Informações Geográficas (SIG) em investigações históricas (2012-2013) e o Projeto Pauliceia 2.0 que exploram as possibilidades das geotecnologias em investigações históricas. Estes estudos foram desenvolvidos pelo Grupo de Pesquisas Hímaco – História, Mapas e Computadores e contaram com a participação e o apoio do Núcleo de Acervo Cartográfico do Arquivo Público do Estado de São Paulo (Apesp).

Além dos projetos anteriores, houve também o desenvolvimento do Projeto de avaliação e edição de verbetes da Enciclopédia Virtual Wikipédia foi realizado durante as disciplinas de História Moderna I e II na UNIFESP, na qual se propôs a alteração dos verbetes na Wikipédia. Este projeto recebeu suporte da equipe brasileira ligada à Wikipédia Foundation que ofereceu palestras e orientou sobre os primeiros passos para a edição do conteúdo (Ferla, Lima & Feitler, 2020).

Com o objetivo de incluir o projeto Pauliceia 2.0 no âmbito da ciência aberta o projeto foi apresentado em 2017 no auditório do Arquivo do Estado de São Paulo e propôs-se a discutir, receber críticas e sugestões. Além disso, aproveitou-se este momento para solicitar material empírico para dar suporte aos testes desenvolvidos dos códigos computacionais. Em outubro de 2018, uma nova apresentação foi realizada para apresentação da versão beta da plataforma na qual se convidou a comunidade para ajudar nos testes (Ferla, Lima & Feitler, 2020).

No mesmo ano, Ferla et al. (2020) apresentam os resultados parciais do projeto Pauliceia 2.0: mapeamento colaborativo da história de São Paulo (1870-1940) que é uma iniciativa aberta e colaborativa que pretende mostrar as transformações urbano-industriais ocorridas na cidade de São Paulo. O projeto propõe a espacialização dos dados por meio de visualizações cartográficas a partir de dados geográficos do passado. Dessa forma, foi estruturado em banco de dados com endereços conhecidos que foram levantados e organizados pro historiadores, enquanto a equipe computacional desenvolvia algoritmo para a ferramenta.

Segundo os autores projetos dessa natureza não apresentam resultados definitivos, pois são constantemente reavaliados e aprimorados. Além disso, projetos que se propõe abertos, no qual a participação de qualquer interessado seja para consulta ou para a alimentação dos dados, correm o risco de perda de qualidade nos dados inseridos. Portanto, é fundamental que o usuário que alimenta os dados seja responsável pela confiabilidade das informações inseridas e que a própria plataforma funcione como um controle da qualidade do que foi alimentado nela (Ferla et al., 2020).

Albagli, Py e Iwana (2020) apresentam o projeto de pesquisa-ação Ciência Aberta Ubatuba no qual se desenvolveu o protótipo chamado de LindaGeo - Litoral Norte Dados Abertos Geoespaciais, com o objetivo de testar a geovisualização de dados abertos gerados naquele território. Segundo os autores, a experimentação ocorreu em 2014, quando houve a revisão do Zoneamento Ecológico Econômico do Litoral Norte de São Paulo (ZEE-LN) com o objetivo de “promover o ordenamento territorial e disciplinar os usos dos recursos naturais da região.” (Albagli; Py & Iwana, 2020).

Os autores observaram desconforto por parte das comunidades locais, pois muitas consultas não haviam sido conduzidas de modo participativo e informativo. O LindaGeo procurou suprir essa lacuna ao "mobilizar, envolver e trocar experiências, demandas e expectativas entre diferentes grupos locais" (Albagli; Py & Iwana, 2020). Contudo, para que o uso dos softwares utilizados fosse efetivo foi necessário desenvolver capacitações, tutorias e a abordagem de forma lúdica, para que os grupos locais pudessem fazer uso dessas plataformas. Apesar disso, os participantes encontraram dificuldade.

Os autores Baum (2017), Steinerova (2018), Knöchelmann (2019), Hocker, Schindler e Rittberger (2020) e Cantón (2021) que abordam as Humanidades Digitais e a Ciência Aberta, mas não utilizam projetos para o embasamento serão abordados a partir desse ponto.

Baum (2017) em seu artigo *Digital gap or Digital turn?* aborda os estudos literários e a era digital. Para a autora o digital trouxe uma série de vantagens, pois nos ambientes digitais os códigos, os *scripts* e as decisões de anotações podem ser divulgadas e até mesmo, discutidas e revisadas de maneira facilitada por meio do versionamento. Contudo, a autora também traz a tona questões cruciais como: a lacuna entre os

países de primeiro e terceiro mundo provocada pelas tecnologias, além das questões socioeconômicas, tais como etnia, gênero, nacionalidade e educação, que incluem acesso a alfabetização digital.

No contexto da ciência aberta, Baum (2017) afirma que a capacidade de conectar a própria pesquisa não prova que tudo pressupõe processos digitais, ou que tudo está subordinado. Principalmente, porque há algumas questões impostas pelo digital que são difíceis de refletir no ambiente analógico devido a sua alta complexidade técnica, que requer uma expertise digital e um trabalho colaborativo nesse setor. Segundo a autora o termo Open Science agrupa estratégias e processos que tem como objetivo a oportunidade de digitalização para tornar os componentes do processo científico acessíveis e reutilizáveis, ou seja, trazer novas oportunidades para a ciências, a sociedade e os negócios.

Baum (2017) observa que os estudos literários são um grande desafio para a ciência aberta, pois é um campo vasto e heterogêneo com uma ampla variedade de autores e uma série de pesquisadores individuais, grupos de trabalho, institutos, agências de financiamento, editoras, provedores de internet e público experiente em rede. Assim, deve-se permitir o acesso aberto à literatura científica especializada, garantindo que ela possa ser referenciada e citada, além de promover novas formas de comunicações especializada e o reconhecimento de pesquisas colaborativas e realizações de publicações.

No artigo *Perceptions of the information environment by researchers: a qualitative study* (2018), Jela Steinerova procura compreender como a pesquisa e as práticas acadêmicas representam um desafio para a melhoria dos serviços de informação e das infraestruturas de informação. Em sua pesquisa a autora buscou identificar as percepções e usos de recursos de acesso aberto e ciência aberta por pesquisadores. Steinerova identificou que os pesquisadores consideram vantajoso o acesso aberto, principalmente pelo aumento de citações e velocidade de publicação, mas também expressam preocupações com as influências comerciais e as avaliações das publicações digitais (Steinerova, 2018).

Steinerova (2018) também observou que muitos pesquisadores concordam com as políticas europeias de acesso aberto, ao passo que outros temem a qualidade inferior da publicação digital. A transparência e o acesso aberto foram identificados como fatores de ciência aberta, bem como a participação, a colaboração, a rede de pares e o compartilhamento de informações. A determinação tecnológica foi identificada nas ciências de big data, como astrofísica, física, genética e outras. Nas Humanidades, foi observado o desenvolvimento de bibliotecas e arquivos digitais no patrimônio cultural. Além disso, foram mencionados outros fatores de ciência aberta, tais como: políticas, avaliação de resultados, acesso a dados e publicação.

Ainda segundo a pesquisa de Steinerova (2018), as redes sociais de pesquisadores são utilizadas para compartilhamento de dados, informações e publicações. A ciência aberta foi percebida como uma vantagem, principalmente as fontes em acesso aberto, além da cooperação interdisciplinar e a publicidade, porém também foram observadas lacunas na coordenação da ciência aberta e no acesso a publicações e dados, bem como preocupação quanto as influências comerciais e o acesso ao financiamento.

Marcel Knöchelmann, em seu artigo *Open science in the humanities, or: Open humanities?* (2019) afirma que a ciência aberta tem deficiências para abordar as ciências humanas, que por isso é necessário pensar um discurso sobre as humanidades abertas. Para o autor os argumentos para a necessidade desse discurso

são: as humanidades são um subproduto da ciência aberta, pois não tem um discurso próprio; a fragmentação dos discursos sobre as práticas abertas nas humanidades requer uma unificação desses discursos e principalmente, as inadequações das práticas atuais de comunicação científica, uma vez que há diferenças nas práticas de comunicação das disciplinas científicas e humanas (Knöchelmann, 2019).

Segundo Knöchelmann (2019) o termo humanidades abertas já foi utilizado anteriormente, mas isso não significa que exista um discurso de humanidades abertas. Para o autor a importância de um discurso de humanidades aberto reúne pensar que as humanidades precisam estar abertas à transferência do conhecimento interdisciplinar, principalmente no que tange as humanidades digitais; as humanidades também precisam de um espaço transdisciplinar no qual possam moldar seu futuro digital e aberto, trabalhando na abertura de suas práticas, seus problemas e implementações. Outra questão apontada por Knöchelmann (2019) diz respeito ao uso de licenças de direitos autorais, uma vez que elas são importantes para o entendimento progressivo da autoria.

Hocker, Schindler e Rittberger (2020) propõe um projeto de design participativo para criar conceituações compartilhadas para ontologias. Segundo os autores existem diferentes abordagens para projetar ontologias, porém, normalmente, o usuário participa do processo de criação ou, quando participa, é apenas para avaliar algo que já foi projetado. Os autores pretendem ir além dessas práticas, propondo um processo de design de ontologias utilizando o método de design participativo ou codesign, que garanta maior qualidade e conhecimento sobre os produtos e sistemas construídos, que facilite a sua aceitação e reduza a resistência a mudança e que permita que as pessoas participem da tomada de decisão, aumentando a democracia no local de trabalho.

Ao utilizarem conhecimentos das humanidades digitais e da ciência aberta, os autores promovem a inserção dos usuários desde o estágio inicial do design, permitindo o compartilhamento de conhecimentos, a democratização dos processos e a possibilidade de participar da comunidade com suas habilidades e opiniões. Por fim, observaram e concluíram que o design participativo trouxe o comprometimento dos usuários, a aprendizagem mútua, a qualidade do feedback durante o processo e a qualidade da ontologia geral (Hocker, Schindler, & Rittberger, 2020).

O estudo proposto por Cantón (2021) buscou identificar o acesso às principais revistas de Humanidades, porém segundo a própria autora o estudo apresenta limitações, pois inclui apenas revistas que publicam em inglês e espanhol. A pesquisa identificou 88 periódicos de Humanidades Digitais, destes, 21 encontram-se encerrados. Para delimitar o corpus da pesquisa elencou-se os seguintes critérios: ser um periódico acadêmico, publicar mais de 50% dos seus artigos sob a temática das Humanidades Digitais, estar ativa a pelo menos 5 anos e ter publicação a partir de 2010 e ser reconhecida pelos pesquisadores da área.

A pesquisa identificou 38 revistas de Humanidades Digitais que publicam majoritariamente em acesso aberto, utilizando a via diamante. Constatou-se que mesmo os editores comerciais publicam em acesso aberto, porém seria desejável que os encargos fossem reduzidos e o acesso às publicações, mais acessíveis, para que os resultados das pesquisas fossem divulgados e reaproveitados, que é um dos princípios das Humanidades Digitais (Cantón, 2021) e da Ciência Aberta.

Ênfase à Ciência Aberta

Alguns artigos do corpus enfatizam mais que os outros alguns aspectos característicos da ciência aberta. Estes aspectos serão abordados nesta sessão.

Basset, Di Giorgio e Ronzino (2017) enfatizam que por ser um projeto do Programa Horizonte 2020 a gestão de dados deve preocupar-se em atender aos princípios FAIR (*Findable, Accessible, Interoperable, Reusable*), ou seja, os dados devem ser localizáveis, acessíveis, interoperáveis e reutilizáveis. Assim, os projetos financiados pelo Programa devem implementar um Plano de Gerenciamento de Dados (PGD) com o objetivo de melhorar e maximizar o acesso e a reutilização dos dados de pesquisa gerados pelo projeto. O que reforça as ações da União Europeia para a circulação do conhecimento.

Os autores afirmam que o modelo de PGD PARTHENOS está dividido em três níveis: primeiro nível inclui um conjunto de requisitos gerais essenciais, independente da disciplina; o segundo nível inclui requisitos específicos e o terceiro nível é baseado em projeto. Para auxiliar o preenchimento do PGD será fornecido um conjunto de instruções para as disciplinas específicas. Segundo os autores, o modelo PGD PARTHENOS permite que os pesquisadores acessem, minerem, explorem, reproduzam e divulguem livremente seus dados e identifiquem as ferramentas necessárias para utilizar os dados brutos e validar os dados da pesquisa, ou para que forneçam suas próprias ferramentas, dando um passo significativo para a realização da ciência aberta (Basset, Di Giorgio, & Ronzino, 2017)

Wuttke, Spiecker e Neuroth (2019) enfatizam aos princípios FAIR e os valores da Ciência Aberta, aumentando a cooperação entre as infraestruturas de pesquisa já existentes, estimulando o intercâmbio de padrões tanto técnicos quanto semânticos, e ampliando as fronteiras disciplinares, fazendo com que as discussões saiam dos muros das comunidades científicas e passem a envolver representantes de centros de dados, instituições de memória e associações de pesquisa. Segundo os autores, essas ações são uma importante contribuição para a *European Open Science Cloud* (EOSC), que se propõe a ser um ambiente no qual os pesquisadores possam armazenar, analisar e reutilizar dados para fins de pesquisa, educação, entre outros.

Segundo Hagmann (2018) o PHAIDRA é baseado no *Fedora Commons Repository* e apresenta informações sobre autoria, licenciamento, quadro histórico da pesquisa entre outras. Além de incluir informações sobre o versionamento, classificação e categorização mais detalhadas. Todos os objetos e coleções tem identificadores persistentes (link permanente, *handle* e DOI), facilitando a citação. Todas essas ações permitem a reutilização dos dados. Além disso, sempre que possível os registros de dados recebem uma licença Creative Commons Atribuição 4.0 Internacional (CC BY 4.0) com relação ao acesso livre e aberto.

Para Lathi et al. (2019) o LOD representa um passo crucial para aproveitar ao máximo os recursos digitais por meio da integração de fontes da web e metadados abertos e reutilizáveis.

De acordo com Anglada (2019) o movimento da Ciência Aberta faz com que a ciência seja mais rápida, precisa e reutilizável. Para o autor há três critérios para se fazer a Ciência Aberta, ela deve ser aberta, colaborativa e feita com e para a sociedade. O que se observa, segundo o autor, é que o movimento da Ciência Aberta e das Humanidades Digitais são um reflexo das profundas mudanças nos processos de investigação e pesquisa. Assim, não se trata de fazer mais divulgação científica, mas de estabelecer novas

relações entre as pessoas e a ciência. Os cidadãos não só podem interessar-se pela ciência, como podem contribuir com sua participação e suas contribuições.

Stiegler e Klug (2019) afirmam que o projeto KONDE possui uma orientação consistente para a Ciência Aberta (*Open Access, Open Data e Open Source*). Além disso, o projeto garante o acesso a longo prazo e de forma gratuita aos dados de pesquisa e permite que leigos se envolvam ativamente, se necessário, e cientificamente, se indicado, no sentido da abordagem da ciência cidadã.

Segundo Rollo (2020) a ciência aberta tem um enorme potencial de transformação, principalmente no que tange a atuação das Humanidades Digitais, uma vez que alarga as interconexões entre as humanidades e a sociedade. Segundo a autora, a ciência aberta não se concentra apenas no acesso aberto, na inovação aberta, na ciência cidadã, mas contempla também os desafios do arquivamento e armazenamento, preservação e curadoria dos dados e informação digital que é produzida em larga escala pela sociedade.

Desafios para a Ciência Aberta

Já em 2014, Wells et al. enfatizaram que a questão financeira é uma fonte de preocupação, pois o financiamento de bens públicos, como dados abertos normalmente requer o apoio público ou filantrópico. Além disso, os autores ressaltam que a sustentabilidade financeira do projeto requer que os legisladores compreendam que os bancos de dados são um patrimônio nacional da América.

Outro desafio mencionado por Wuttke, Spiecker e Neuroth (2019) é o grande volume de dados, ferramentas e métodos digitais que ainda não foram tratados de forma adequada pelas comunidades científicas. Além disso, os autores enfatizaram que as comunidades e os participantes do projeto apresentam déficits a respeito dos princípios FAIR e do conceito EOSC.

Hagmann (2018) descreve como desafio a grande diversidade de dados e suportes obsoletos. Dessa forma, o desafio é tornar esses dados arqueológicos disponíveis de forma sustentável, por meio de um arquivamento eficaz a longo prazo. Para isso é necessário que os dados sejam salvos em formatos adequados para arquivamento, o que permite o desenvolvimento de pesquisas futuras, bem como a preservação destes dados.

Lathi et al. (2019) enfatizam que a falta de disponibilidade de dados abertos constitui um grande gargalo para o desenvolvimento transparente e colaborativo da ciência de dados bibliográficos. Além disso, criticam, de certa forma, o termo ciência aberta, pois para os autores essa terminologia não se preocupa com as humanidades, mas com um agrupamento de disciplinas científicas.

Albagli, Py e Iwana (2020) observaram que a cartografia social, associada a ciência aberta e cidadã podem contribuir não apenas com a representação do território, mas permitir que grupos e populações locais possam integrar seus conhecimentos a esse tipo de projeto. Contudo, é necessário que os grupo e populações locais sejam incluídas nas discussões e possam participar ativamente dos processos decisórios.

CONCLUSIONES

Apesar da Declaração de Budapeste datar de 2002 e do movimento da ciência aberta ter sido impulsionado por ele (Santos, Almeida & Henning, 2017), encontramos publicações que relacionam as Humanidades Digitais a ciência aberta apenas a partir de 2014 e com relação ao volume de publicações apenas em 2019 e 2020 há um número de publicações expressivas sobre a temática.

Esperava-se que os mineradores de textos evidenciassem a relação entre as humanidades digitais e a ciência aberta, porém isso não foi observado. No corpus analisado não ficou clara a relação entre as humanidades e o digital, nem entre os termos ciência e aberta. As nuvens de palavras tornaram possível a observação de termos mais recorrentes nos resumos. Este problema pode ter ocorrido pela utilização do resumo para confecção do corpus, uma vez que os resumos apresentam uma limitação de palavras para sua elaboração.

Com relação aos achados sobre a ciência aberta observa-se a importância do Plano de Gerenciamento de Dados (PGD) bem elaborado. Para que um PGD seja bem elaborado é fundamental que existam documentos que orientem o pesquisador quanto a elaboração de um PGD para cada área do conhecimento, uma vez que cada área apresenta diferentes especificidades.

A padronização dos metadados, o uso de softwares que permitam o versionamento dos dados da pesquisa, a adoção de links permanentes, a atribuição de licenças *Creative Commons* e, principalmente, pensar que os dados gerados devem atender aos princípios FAIR, ou seja, os dados precisam ser localizáveis, acessíveis, interoperáveis e reutilizáveis, facilitam a ciência aberta.

Além disso, a criação e o desenvolvimento de materiais de comunicação, a oferta de treinamentos que visem a capacitação dos pesquisadores e cidadãos, bem como a colaboração de profissionais de diferentes instituições, com diferentes formações aliada a participação de cidadãos, cria o ambiente propício para a valorização da ciência aberta.

Essa valorização se estende aos espaços que este patrimônio físico e digital ocupa, ou seja, amplia a importância e dá visibilidades para os espaços de guarda e preservação da memória como bibliotecas, arquivos e museus. Espaços de educação, de guarda patrimonial, de democratização, de recolha, de partilha e de preservação da memória, espaços essenciais para as Humanidades e para o avanço da ciência.

REFERENCIAS BIBLIOGRÁFICAS

Albagli, S. (2015). Ciência aberta em questão. In Albagli, S., Maciel, M.L., & Abdo, A.H. *Ciência aberta, questões abertas*. Brasília, IBICT.

Albagli, S., Py, H., & Iwana, A.Y. (2020). Geovisualização de dados e ciência aberta e cidadã - a experiência da Plataforma LindaGeo. *DHQ - Digital Humanities Quarterly*, 14(2), 2020. <http://www.digitalhumanities.org/dhq/vol/14/2/000452/000452.html>

- Almeida, M.A.de, & Damian, I.P.M. (2017). Humanidades digitais: um campo praxicológico para mediações e políticas culturais. In *XVI Encontro Nacional de Pesquisa em Ciência da Informação*. João Pessoa, PB.
- Anglada, L.M. (2019). Muchos cambios y algunas certezas para las bibliotecas de investigación, especializadas y centros. *El profesional de la informacion*, 28(1).
- Ávila Araújo, C. A. (2021). Development of an Iberoamerican informational thinking. *Iberoamerican Journal of Science Measurement and Communication*, 1(2), 1-3. <https://doi.org/10.47909/ijsmc.87>
- Baum, C. (2017). Digital gap oder digital turn? Literaturwissenschaft und das digitale Zeitalter. *Zeitschrift für Germanistik*, (27), 316-328.
- Bassett, S., Di Giorgio, S., & Ronzino, P. (2017). A data management plan for digital humanities: the PARTHENOS Model. *Ercim News*, Special Theme (111).
- Blumesberger, S. (2015). Der welt der metaddaten im universum von repositorien. *Mitteilungen der VÖB*, 68(3/4), 515-528.
- Capes (2020). *Critérios de classificação Qualis – Ensino*.
- Cantón, C.I.M. (2021). Un análisis cuantitativo del acceso abierto en las revistas de Humanidades Digitales. *Revista General de Información y Documentación*, 31(1), 331-348.
- Dalbello, M. (2011). A genealogy of digital humanities. *Journal of documentation*, 67(3), 480-506.
- Ferla, L., Lesser, J., Ferreira, K.R., Miyasaka, C., Atique, F., Musa, D., Britt, A.G., Rogers, T.D., Fook, K.D., & Vijaykumar, N. (2020). Pauliceia 2.0: mapeamento colaborativo da história de São Paulo, 1870-1940. *História, Ciências, Saúde – Manguinhos*, 27(4), 1207-1223.
- Ferla, L.A.C., Lima, L.F.S., & Feitler, B. (2020). Novidades no front: experiências com humanidades digitais em um curso de história na periferia da grande São Paulo. *Estudos Históricos*, 33(69), 111-132.
- Gtech.Edu. *Mineração de texto educacional*. Recuperado de http://sobek.ufrgs.br/uploads/sobek_quick_reference_guide_pt.pdf
- Guerreiro, D. (2020). Padre Roberto Busa, S.J. *Bibliotecas e Humanidades Digitais*. <http://https://bdh.hypotheses.org/tag/index-thomisticus>
- Hagmann, D. (2018). Überberlegungen zur nutzung von PHAIDRA als repositorium fü r digitale archäologische daten. *Mitteilungen der VÖB*, 71(1), 53-69.

- Hocker, J., Schindler, C., & Rittberger, M. (2020). Participatory design for ontologies: a case study of an open science ontology for qualitative coding schemas. *Aslib Journal of Information Management*, 72(4), 671-685.
- Larhud. (2020). *Voyant Tools*. http://www.larhud.ibict.br/index.php?title=Voyant_Tools
- Lahti, L., Marjanen, J., Roivainen, H., & Tolonen, M. (2019). Bibliographic data science and the history of the book (c, 1500-1800). *Cataloging & Classification Quarterly*, 57(1), 5-23.
- Londoño, F. (2020). Entrevista com Felipe Londoño [Entrevista concedida à Isabel Jungk]. *TECCOGS - Revista Digital de Tecnologias Cognitivas*, 21, 12-23. doi: <https://doi.org/10.23925/1984-3585.2020i21p12-23>
- Knöchelmann, M. (2019). Open science in the humanities, or: open humanities? *Publications*, 7(65), 1-17.
- Rollo, M.F. (2020). Desafios e responsabilidades das humanidades digitais: preservar a memória, valorizar o patrimônio, promover e disseminar o conhecimento. Programa Memória para todos. *Estudos Históricos*, 33(69), 19-44.
- Santos, P.X.dos, Almeida, B.deA., & Henning, P. (Orgs.) (2017). *Livro verde - ciência aberta e dados abertos: mapeamento e análise de políticas, infraestruturas e estratégias em perspectiva nacional e internacional*. Fiocruz, Rio de Janeiro.
- Scopus. (2020). *Sources*. <https://www-scopus.com.ez22.periodicos.capes.gov.br/sources.uri>
- Steinerova, J. (2018). Perceptions of the information environment by researchers: a qualitative study. *Information Research*, 23(4).
- Stiegler, J., & Klug, H.W. (2019). KONDE – Ein netzwerk bringt forschuns – und GLAM-institutionen zusammen. Ein projektbericht. *Mitteilungen der VÖB*, 72(2), 431-439.
- Web of Science Group. *Master journal list*. <https://mjl.clarivate.com/home>
- Wells, J.J., Kansa, E.C., Kansa, S.W., Yerka, S.J., Anderson, D.G., Bissett, T.G., Myers, K.N., & Demuth, R.C. (2014). Web-based discovery and integration of archaeological historic properties inventory data: The Digital Index of North American Archaeology (DINAA). *Literary and Linguistic Computing*, 29(3), 349-360.
- Wuttke, U., Spiecker, C., & Neuroth, H. (2019). PARTHENOS – eine digitale Forschungsinfrastruktur für die Geistes - und Kulturwissenschaften. *Bibliothek – Forschung und Praxis*, 43(1), 11-20.