

# A congruência intelectual da ciência da informação nas Humanidades Digitais

## The intellectual congruence of information science in the Digital Humanities

Renan Castro, Fundação Getulio Vargas, Brasil, <https://orcid.org/0000-0001-5605-6212>

**Resumo:** Investiga uma noção sensível daquilo que pode revelar a influência da literatura acadêmica na composição das humanidades digitais. Elabora um mapeamento do campo através da análise de métricas da bibliometria. Implementa uma leitura desses dados a partir da utilização do software VOSviewer para construção de um diagnóstico do escopo das humanidades digitais. Aplica uma análise baseada em três pilares dos estudos bibliométricos: autores, periódicos e palavras-chave dos artigos científicos recuperados. Esboça uma análise de escopo do domínio das humanidades digitais a partir da investigação bibliométrica. Clarifica a constituição de um domínio incipiente e complexo através da análise das características impressas na literatura e somente através dela disponível. Persegue uma categorização capaz de não apenas revelar detalhes de seu escopo, mas também fornecer elementos indicativos de quais temáticas influenciam a área. Indica assim a influência da Ciência da Informação nas humanidades digitais e visualiza com clareza que termos se apresentam com maior recorrência.

**Palavras-chave:** humanidades digitais; ciência da informação; análise bibliométrica; análise de domínio.

**Abstract:** The paper investigates a sensitive notion of what may reveal the influence of academic literature production on the composition of the digital humanities. It prepares a field mapping through the analysis of bibliometric metrics. It implements a reading of these data using the VOSviewer software to build a diagnosis of the scope of digital humanities. It applies an analysis based on three pillars of bibliometric studies: authors, journals and keywords of retrieved academic articles. It outlines a scope analysis of the digital humanities domain based on bibliometric research. It clarifies the constitution of an incipient and complex domain through the analysis of characteristics printed in literature and only available through it. It pursues a categorization capable of not only revealing details of its scope, but also providing elements indicative of which themes influence the area. Thus, it indicates the influence of Information Science on the digital humanities and clearly visualizes which terms are most recurrent.

**Keywords:** digital humanities; information science; bibliometric analysis; domain analysis.

## 1. Introdução

No contexto de produção exponencial de informação, seguido de perto pela profusão de ferramentas computacionais aplicadas às metodologias de pesquisa num ambiente alheio às funções dessa ordem, se acha espaço para novas reflexões. A ideia de sociedade da informação articula esse movimento dentro de uma lógica na qual o processo não se caracteriza pelo mero ciclo de realimentação cumulativo entre as inovações e seus usos, como parece ser consagrado na sociedade em rede. Pela sociedade da informação, segundo Marques & Pinheiro (2013, p. 117) "cria-se a necessidade do estabelecimento de novas políticas, que atendam a tais transformações". Para as autoras destacam-se neste cenário discussões relativas ao acesso à informação e aos atores responsáveis pela função de democratizá-la. Com efeito, vemos, a partir de então, um direcionamento dessas reflexões voltando-se às problemáticas do campo informacional ou este se apropriando daquelas.

Foi no campo informacional que essas mudanças se fizeram sentir ainda com maior voracidade, sobretudo diante da constatação da tecnologia pautar autoritariamente a realidade de instituições de informação como bibliotecas, centros de documentação e afins. Esse processo, aliado à explosão informacional<sup>3</sup>, que Varela (2005, p. 2) pontua como o movimento em que destacam-se "a abordagem da realidade complexa através de ferramentas cognitivas; a construção de representações mentais; o desenvolvimento de competências e habilidades; e a reflexão em torno da apreensão e da compreensão da informação pelo sujeito", que na nossa avaliação atribuiu relevância ao que hoje se desenvolve como campo da Ciência da Informação.

Todavia há uma reflexão específica, oriunda de indivíduos que experimentaram a metodologia tradicional elevada a uma nova percepção, que tem na tecnologia a maior responsável pelo processamento da informação, e que compôs uma comunidade discursiva que vem denominando de humanidades digitais o fenômeno da aplicação de métodos computacionais nas práticas de pesquisa das áreas humanas. Assim, o resultado mais atual desse fenômeno é a expressão mais genuína da nova dimensão metodológica que surge como campo de conhecimento, delineado pelas experimentações praxiológicas e sob o rótulo de humanidades digitais.

Tais características fundamentais nos provocam identificar convergências e serve de mote para buscarmos perseguir relações de similitudes potencialmente contributivas para aquilo que vem atendendo por humanidades digitais. Visando, assim, uma análise perceptiva da congruência de escopos a partir da interpretação do domínio das humanidades digitais e se inspirando, ainda, numa mensuração da influência de outros campos acadêmicos nas HDs. Cremoshaver deliberadas relações entre a Ciência da Informação e aquela que consideramos representar a expressão conceitual desta nova dimensão de interpretação da aplicação tecnológica como instrumento metodológico para aproveitamento da informação, ou seja, as humanidades digitais.

Assim a proposta deste estudo é identificar nas humanidades digitais, produto e sintoma das transformações tecnológicas, aquilo que se identifica com interdisciplinaridade. Essa característica é determinante para motivação de rastrear alguma relação de proximidade que nos corroborasse a ideia de que, além da mera relação interdisciplinar, revelasse, também, uma justaposição. Nesse contexto, o esforço é atribuir o papel da Ciência da Informação,

identificando-o enquanto congruência de escopo capaz de modular e influenciar os debates do novo campo.

## **2. A influência cognitiva como traçado de uma justaposição entre ciência da informação e humanidades digitais**

O que outrora centrava nos estudos bibliométricos a base para avaliação da ciência e da produção do conhecimento científico, contemporaneamente gestou um aprimoramento das ações inspiradas por essa técnica. As medidas basicamente estatísticas de antes, atualmente estão associadas a técnicas que permitem entre diversas formas de análise, a visualização das redes de conteúdos que possibilitam a materialização de uma emaranhada estrutura das relações entre métricas da bibliometria.

Assim, nesse contexto, ganha corpo a análise de citações que fundamenta a construção de uma rede cognitiva entre indivíduos. De Bellis (2009) argumenta, enquanto fundamento para estabelecimento de ligações entre autores e/ou trabalhos, que a revelação da estrutura de um campo estaria aplicada ao que se refere como 'indicador de proximidade cognitiva'. Vemos, portanto, nessa afirmação, a materialização do que Zhao & Strotman (2015) buscam mapear enquanto estrutura intelectual.

Para Garfield (1983, p. 81, tradução nossa) o uso da "análise de citações em pesquisas sobre a estrutura da ciência é baseado em um modelo do processo científico". Neste modelo, o trabalho científico é representado pelos artigos escritos e publicados para reportá-lo, além das relações entre trabalhos que são representadas pelas referências nos artigos. Tanto Garfield como Price, autores expoentes na construção desse modelo de estudo do conhecimento, adotaram a ideia de que os artigos científicos compreendem o principal produto do trabalho de um cientista, servindo, portanto para análise do conhecimento. Para Leydesdorff (1998, p. 5, tradução nossa) esse foi o contexto de "surgimento de conceitos como 'mapeamento da ciência' através da introdução dos estudos da dinâmica das redes de citação".

## **3. A visualização bibliométrica das redes cognitivas como metodologia**

Assim sendo, a análise proposta por esta pesquisa está baseada na investigação da literatura produzida e indexada por 'humanidades digitais' nas bases de dados consolidadas e consideradas instrumentos agregadores de literatura científica de natureza multidisciplinar. Dessa forma, a estratégia metodológica foi revisar a literatura sobre o tema na base Scopus e processá-la via software de visualização de informações bibliográficas VOSviewer.

Dessa forma aplicou-se uma expressão de busca para contemplar o maior alcance possível na literatura, recuperando assim a publicação sobre *digital humanities* em inglês, francês, espanhol e português para proceder à pesquisa na base de dados eleita para esta revisão. Aplicamos então como padrão o filtro 'Articletitle,abstractandkeywords' ou equivalente que nos reportou a expressão de busca tanto nos títulos quanto resumos, além das próprias palavras-chave dos artigos. Dessa forma a expressão de busca desenhada para esse levantamento pode ser representada por

("Dig ital Humanit \*" OR " humanit \* comput \*" OR " humanidad \* digita \*" OR "digital research\*" OR "digital research\* activit \*" OR "digital methodolog \*")

O primeiro termo foi enriquecido por *string* para abranger variações como, por exemplo, o termo no singular; o segundo termo dá conta da pré-nomenclatura da área, já que a mesma antes de possuir seu nome oficial cunhado apenas em 2010, apresentava nessa expressão seu conceito; o terceiro termo contempla um termo guarda-chuva, apresentado pela TADIRAH, a principal taxonomia da área, como termo apropriado para abarcar todas as práticas das HDs. O último termo visou descrever na prática a principal atividade do campo quando percebida por autores desconhecedores da nomenclatura oficial.

Para processar, com apoio ferramental, a informação extraída, utilizamos os dados obtidos em formatos padrão passíveis de compatibilidade com *softwares* como VOSviewer. Os formatos utilizados foram aqueles de natureza tabular que se aplicam em planilhas estruturadas como Excel. Nesse quesito, portanto, atribuímos preferência ao padrão .CSV. Nós utilizamos, inicialmente, das possibilidades do VOSviewer, que tem o objetivo de servir à estudos de grande parte das métricas da bibliometria de forma automatizada, agregando à elas a possibilidade de visualização de grafos ou mapas de redes, além de um ranqueamento de itens analisados por recorrência e peso.

Nos 1.836 documentos recuperados foram identificados 4.530 autores e destes elegemos analisar aqueles que possuem autoria em pelo menos 3 documentos. Esse universo restringe um grupo de 230 autores que atendem essa exigência. Quando recortados pelos números de citações, na qual elegemos como nota de corte o número correspondente ao dobro dos artigos produzidos por esses autores, chegamos ao grupo de 134 autores.

Tabela 1 – Autores mais citados por força de link. Fonte: elaborado pelo autor.

<b>Author</b>	<b>Documents</b>	<b>Citations</b>	<b>Total link strength</b>
Blanke t.	13	95	35
Hedges m.	10	66	24
Dunn s.	7	40	18
Terras m.	15	202	15
Anderson s.	4	23	13
Conlan o.	10	44	13
Hampson c.	9	43	13
Brown s.	7	25	12
Lawless s.	9	32	11
Bailey e.	5	29	9

A Tabela 1 sintetiza a análise indicando os 10 autores mais relevantes segundo ordem de '*total link strength*'. Esse recorte se mostra suficiente para exibir o grupo seletivo de autores que possuem destacada relevância nessa análise, além dos diversos que dividem uma relevância equiparada entre si. Ainda assim merece destaque a posição em destaque de 'terras m.', (Melissa Terras) ligada academicamente ao campo dos estudos informacionais

(*InformationStudies*) lidere exponencialmente as citações da área, apresentando o maior número de citações todavia sem ocupar o primeiro lugar do *ranking*. A análise da métrica 'citação' nos sinaliza uma tendência considerável de que os autores mais citados constituem o próprio campo das humanidades digitais, tendo suas atuações acadêmicas ligadas a esta temática e nenhuma outra.

A análise observou também os periódicos vetores da literatura das humanidades digitais. Foi identificado um total de 1154 fontes e para estratégia de refino consideramos que o periódico deveria ter publicado pelo menos 5 artigos sobre o assunto. Deste recorte elegemos apenas os periódicos que alcançassem o dobro do número de publicações em citações, ou seja deveriam ser citados 10 vezes no mínimo. Estas exigências compreenderam um grupo de 51 periódicos qualificados após desambiguação.

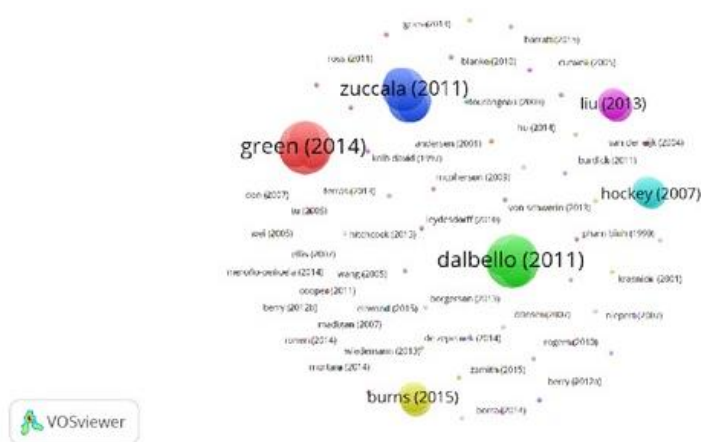
Tabela 2 – Periódicos mais citados por força de links. Fonte: elaborado pelo autor.

Source	Documents	Citations	Total link strength
Journal of library administration	10	209	53
Literary and linguistics computing	56	326	50
Supporting digital humanities for knowledge acquisition in modern libraries	12	12	50
Journal of documentation	9	152	39
Historical social research	16	108	22
Arts and humanities in higher education	11	81	16
Computers and the humanities	20	181	14
Digital scholarship in the humanities	36	77	13
Acm international conference proceeding series	77	78	12
Bijdragen en mededelingen betreffende de geschiedenis der nederlanden	7	38	12

A Tabela 2 explica quais periódicos apresentam-se como mais relevantes quando, de acordo com o recorte criado, alcançam maior índice força de *links*. Além de permitir identificar um grupo mais contundente, restringindo um núcleo mais relevante (3 periódicos com índice 53 a 50), também revela uma distribuição intermediária com índice a partir de <39. Os contornos intelectuais podem ser extraídos se observarmos a área de foco destes periódicos em evidência, surpreende um título devotado aos assuntos de biblioteca/Ciência da Informação estar no topo desta análise. A análise desta métrica (Tabela 2) evidencia uma persuasão da área de bibliotecas (*Library and Information Studies*) como área com periódicos mais citados, caso que não ocorre em proporção semelhante com nenhuma outra área.

Um outro esforço que podemos empreender para auxiliar o mapeamento intelectual das humanidades digitais pode se dar em torno da citação dos documentos. Dependendo de sua área de origem ou seu meio de veiculação, podemos realizar interpretações que sugerem indicações cognitivas. Dos mais de 1.800 itens recuperados sob análise, concentram relevância enquanto influência intelectual nos artigos que citaram um universo de 1.231 que sequer receberam qualquer citação. Visando selecionar a amostra, atribuímos como critério de nota de corte o número de 20 citações: atenderam a esse critério 93 documentos. O maior número de citações recebidas por um documento foi 187, formando um núcleo concentrado com outros 5 documentos mais citados, distribuídos nesse intervalo até o número de 103 como mais baixo de citações nesse grupo. A ver, abaixo, a visualização gráfica desta análise (Figura 1):

Figura 1 – Grafo de documentos com maior número de links entre os mais citados.



Os quatro documentos mais relevantes desta métrica (primeiras quatro células da Tabela 3) são provenientes de periódicos da área de bibliotecas/Ciência da Informação, o primeiro foi publicado pela *Library Quarterly* e trata da facilitação da prática das humanidades digitais por bibliotecários em colaboração em pesquisa e treinamento de codificação de texto (*Facilitating communities of practice in digital humanities: Librarian collaborations for research*). Os próximos três também compreendem publicações de periódicos desse campo como, por exemplo, o segundo mais relevante *The Journal of Documentation* (*A genealogy of digital humanities*), o terceiro da *Journal of the American Society for Information Science and Technology* (*Book reviews in humanities research evaluations*), além do quarto, também do *Journal of the American Society for Information Science and Technology* (*Google book citation for assessing invisible impact*). A análise dos documentos citados reforça ainda mais a tendência de participação da área de Ciência da Informação (*Information Studies*) no contorno intelectual das humanidades digitais.

Tabela 3 – Documentos mais citados por força de links. Fonte: elaborado pelo autor.

Document	Citations	Links
Green (2014)	22	3
Dalbello (2011)	28	3
Zuccala (2011)	29	2

Kousha (2009)	69	2
Burns (2015)	32	1
Burns (2014)	28	1
Anderson (2013)	76	1
Delyser (2013)	37	1
Liu (2013)	42	1
Posner (2013)	53	1

A cocitação também foi objeto desta análise. No caso dos autores, o levantamento possui 38.617 citados nas referências e destes escolhemos analisar aqueles que possuíam pelo menos 38 citações, um número de corte pensado para equivaler a 0,1% do número de autores citados. Destes, apenas 58 autores se enquadram na condição de serem citados pelo menos 38 vezes, uma expressiva parcela de proeminentes indivíduos correspondente a aproximadamente 0,16% do universo de autores citados. Importa atentarmos a esse dado, pois esse grupo representa uma seleção altamente representativa daqueles que se destacam na literatura do campo.

Tabela 4 – Autores mais cocitados por número de citações. Fonte: elaborado pelo autor.

<b>Author</b>	<b>Citations</b>	<b>Total link strength</b>
unsworth, j.	207	2996
moretti, f.	182	1455
kirschenbaum, m.g.	168	2795
drucker, j.	158	2350
mccarty, w.	134	1894
ramsay, s.	131	1684
terras, m.	126	1359
liu, a.	112	1573
svensson, p.	112	1541

A Tabela 4 lista os 10 autores mais citados em cocitações e, neste recorte, podemos verificar que autores de grupos evidentes em outras métricas como 'computação' e 'informação', passam a dividir relevância com autores clássicos como 'moretti, f.' e 'mccarty, w.'. Assim, temos nesta métrica a evidência de que autores como os dois mencionados atuam como expoentes epistemológicos do campo. Por outro lado destacamos a constatação de novos autores despontarem no ranqueamento por cocitação, como é o caso de 'kirschenbaum, m.g.'. Nome ligado aos estudos informacionais, inédito até então, no topo das métricas de autoria, 'kirschenbaum, m.g.' preenche a terceira posição entre os citados em cocitação e chega ao

segundo lugar se observado seu *link* de força. Em paralelo observamos também nomes já destacados reincidirem neste grupo seletivo, promovendo, inevitavelmente, a associação de influência no domínio de suas áreas acadêmicas como já analisadas em métricas anteriores. Servindo de exemplo para ilustrar nossa proposta temos o caso de ' terras, m.'.

Aplicando a análise das citações aos documentos recuperados, temos um universo de 45.668 referências citadas por todos os artigos em análise, utilizamos como linha de corte a função de ter sido citado pelo menos 7 vezes. Esse valor foi trabalhado para que pudéssemos alcançar o mínimo de 10 artigos cocitados. Como o montante de referências citadas é alto, é preciso garantir que nenhuma das referências que compõem o grafo integre o mesmo por acaso.

Tabela 5 – Referências cocitadas por força de links. Fonte: elaborado pelo autor.

<b>Cited reference</b>	<b>Citations</b>	<b>Total link strength</b>
vandegrift, m., varner, s., evolving in common: creating mutually supportive relationships between libraries and the digital humanities (2013) journal of library administration, 53 (1), pp. 67-78	12	18
posner, m., no half measures: overcoming common challenges to doing digital humanities in the library (2013) journal of library administration, 53 (1), pp. 43-52	10	17
sula, c.a., digital humanities and libraries: a conceptual model (2013) journal of library administration, 53 (1), pp. 10-26	10	14
vinopal, j., mccormick, m., supporting digital scholarship in research libraries: scalability and sustainability (2013) journal of library administration, 53 (1), pp. 27-42	9	10
ramsay, s., (2011) reading machines: toward an algorithmic criticism, , urbana: university of illinois press	10	5
gold, m.k., (2012) debates in the digital humanities, , minneapolis: university of minnesota press	9	5
moretti, f., conjectures on world literature (2000) new left review, 1, pp. 54-68	11	3
schreibman, s., siemens, r., unsworth, j., (2004) a companion to digital humanities, , oxford: blackwell	9	2
busa, r., the annals of humanities computing: the index thomisticus (1980) computers and the humanities, 14, pp. 83-90	7	2
manovich, l., (2013) software takes command, , new york: bloomsbury academic	8	1



Pode-se verificar na Tabela 5 que o principal aglomerado (primeiras quatro células) é, em totalidade, o responsável por deter a maior força de *link*, ou seja, todos seus itens possuem as maiores relevâncias nas cocitações. Todos esses itens tratam, sem exceção, da temática 'bibliotecas' e se sobressaltam por serem todos provenientes do '*journal of library administration*'. A análise dessa métrica expõe com ineditismo a predominância absoluta de uma área numa métrica, ou pelo menos, não deixa margem de dúvida sobre essa interpretação. Se somarmos a essa leitura o fato da conexão deste *cluster* com aquele que apresenta duas obras centrais para área, podemos inferir que a área de bibliotecas e informação tem um papel central nos debates em humanidades digitais, contribuindo para delinear e desenvolver o campo a partir de suas influências.

Seguindo ainda nas análises de cocitações, é possível aplicá-la ao corpus de 28.740 fontes referenciadas, e para possibilitar a observação, refinamos a linha de corte entre os que possuíam pelo menos 85 cocitações. Nessa configuração o número de fontes selecionadas atinge as 10 principais fontes da literatura e viabiliza alguma interpretação mais acurada.

Tabela 6 – Principais fontes cocitadas. Fonte: elaborado pelo autor.

Source	Citations	Total link strength
Literary and linguistic computing	342	1240
Digital humanities quarterly	204	751
A companion to digital humanities	186	895
Computers and the humanities	177	745
Debates in the digital humanities	162	573
Journal of the American society for information science and technology	125	782
Science	125	161
Scientometrics	117	664
D-lib magazine	104	209
Journal of library administration	90	424

Mesmo não integrando o topo desta métrica, a área respectiva aos estudos informacionais segue presente nesse recorte mais significativo: única área acadêmica a possuir essa presença ('*d-lib magazine*' no *cluster* 2 e '*journal of the american society for information science and technology*' no *cluster* 3). O recorte ainda exhibe um dado passível de interpretação no qual o *cluster* 3 pode ser considerado como um agrupamento da área dos estudos informacionais

por compreender-se pelo periódico '*scientometrics*'. Configurando assim mais um fator de destaque da área informacional entre as demais por ser esta uma condição ímpar.

Findando a análise dos resultados desse estudo, considera-se de fundamental uma complementação dessa análise bibliométrica, a partir da interpretação das palavras-chave da literatura reunida por este estudo. Dessa forma analisamos nesta pesquisa um total de mais de 7.000 palavras-chave responsáveis por classificar os documentos. Em busca de produzir uma abordagem com alto nível de relevância, elegemos enquanto fator de corte a reincidência de 50 vezes, construindo assim um núcleo de 11 palavras-chave mais importantes nesta revisão de literatura. A nota de corte evidencia a forte variedade de termos, pois apenas 11 termos dentre os mais de 7.000 se repetem uma quantidade de vezes que não pode ser considerada alta.

Table 7 – Palavras-chave 'totais' mais recorrentes (49 principais). Fonte: elaborado pelo autor.

<b>Keyword</b>	<b>Occurrences</b>	<b>Total link strength</b>
digital humanities	643	912
humanities computing	314	599
digital libraries	117	283
research	99	273
education	103	235
metadata	73	183
semantics	55	161
humanities research	53	159
visualization	67	153
history	64	148
natural language processing systems	35	134
data mining	41	129
semantic web	42	127
cultural heritages	34	119
information systems	33	117
social sciences	36	113
information retrieval	34	111
linked data	33	111
computational linguistics	34	107

humanities	54	102
social networking (online)	34	100
data handling	24	98
ontology	32	98
world wide web	32	95
artificial intelligence	38	90
big data	37	90
learning systems	26	81
ecosystems	28	78
behavioral research	24	77
design	29	76
engineering education	30	75
linguistics	25	72
cultural heritage	24	71
computer science	26	70
arts computing	25	69
digital collections	23	69
human computer interaction	28	68
internet	32	68
text mining	24	66
information technology	32	64
research infrastructure	25	62
information management	23	61
information science	24	61
virtual reality	30	55
algorithms	24	54
digital researches	30	50
collaboration	24	46
digitization	24	30

O estudo sobre as palavras-chave nos permite observar uma variedade de termos que remetem às temáticas reveladoras sobre o campo das humanidades digitais. Podemos usá-las para, além de um categorização que abranja áreas/campos, práticas/métodos e objetos/alvos, que nos permita compreender uma espécie de representação temática-conceitual do campo, também conseguir vislumbrar um cenário macro da área como defendem Cheng; Xiao (2016). Nesse sentido podemos categorizar enquanto áreas: *education, semantics, history, social sciences, cultural heritage, information science, linguistics e computer science*. Da mesma maneira enquanto práticas de pesquisa podemos identificar: *humanities research, natural language processing systems, data mining, data handling, artificial intelligence, learning systems, engineering education, text mining, virtual reality, digital researches, collaboration e digitization*. Para objetos das humanidades digitais identificamos nesta categoria as palavras-chave: *digital libraries, metadata, semantics, visualization, information systems, information retrieval, linked data, computational linguistics, ontology, artificial intelligence, digital collections, internet, information technology, research infrastructure, information management, algorithms, digital researches*.

Essas categorias sugeridas servem à compreensão de que as humanidades digitais compreendem-se, segundo análise das palavras-chave mais recorrentes na ordem das 50 principais, pelas áreas acadêmicas da Educação, História, Ciências Sociais, Ciência da Informação, Linguística e Ciência da Computação. Enquanto que suas práticas mais efusivas, considerando esse cenário, se apresentam o processamento de linguagem natural, mineração de dados, realidade virtual, pesquisas em meio digital, métodos colaborativos e digitalização. Na categoria dos objetos para as *digital humanities* podemos identificar as bibliotecas digitais, metadados, visualização, sistemas de informação, recuperação da informação, dados abertos interoperáveis, ontologias de representação da informação, inteligência artificial, coleções digitais, tecnologia da informação, infraestrutura de pesquisa, algoritmos e pesquisas digitais.

Assim sendo consideramos a área de Ciência da Informação aquela com maior efusividade nesta análise não apenas por constar nesta categoria, mas principalmente por estar relacionada concomitantemente aos objetos (bibliotecas digitais, recuperação da informação, metadados e ontologias) e também com as práticas (pesquisas em meio digital, digitalização). A eloquência dessa área se manifesta com relevância nas três categorias que sugerimos, o que não ocorre, em respectiva clareza, com nenhuma das outras áreas identificadas nesta categoria.

#### **4. Considerações finais**

No que tange a análise de citação de autores, periódicos e artigos/documentos podemos considerar que as autoridades mais destacadas, que identificamos uma tendência considerável de que os autores mais citados constituem genuinamente o próprio campo das humanidades digitais, tendo suas atuações acadêmicas ligadas a esta temática e nenhuma outra. Todavia, destacou-se que uma autora ligada academicamente ao campo dos estudos informacionais (*Information Studies*) liderando exponencialmente as citações da área. De toda a forma evidenciamos a participação de autores que não formalmente ligados às *digital humanities*, como oriundos da Ciência da Informação e Computação.

Ao que cabe a citação de periódicos científicos, comprovamos uma persuasão da área de bibliotecas e informação (*Library and Information Studies*) como área com periódicos mais citados, caso que não ocorre em proporção semelhante com nenhuma outra área. Todavia a dispersão disciplinar tenha sido corroborada, foi possível identificar algumas áreas com presença declarada como no caso de títulos das áreas da Computação, Linguística e História. Contudo se detecta apenas uma com participação destacada como o caso da área de bibliotecas e Ciência da Informação (*Library and Information Studies*).

Quando observada a análise dos documentos citados, reforçamos ainda mais a tendência de participação da área de Ciência da Informação (*Information Studies*) no contorno intelectual das humanidades digitais. Apesar da constatação de uma dispersão multidisciplinar ser a regra, percebemos no contexto destas análises, que das áreas que mais se evidenciam com proeminência é a área dos estudos informacionais. Vimos ainda a participação de outras áreas em interação como Geografia, Linguística além de um movimento de consolidação de uma parcela de autores integrando um campo específico de *digital humanities*. Por fim, percebemos a constituição de um núcleo duro da Computação recorrentemente presente, porém, isolado e predominantemente sem interligação com as demais áreas.

A análise de coautorias seguiu comprovando a setorização dos núcleos de autores da produção acadêmica nas humanidades digitais, uma vez que nele é possível identificar a incidência de núcleos de áreas acadêmicas distintas formando redes isoladas das outras, embora com importância equilibrada. De fato a análise de coautoria reflete, naturalmente, a relevância de obras referenciais para uma área – recente e com alto nível de necessidade de estabelecimento epistemológico. Todavia consideramos a diversificação de coautorias como um fator relevante no quesito influência, pois está posto que indivíduos da área de *Information Studies* possuam maior influência do que os atores do *cluster* principal (maior número de autores em coautoria e que podem ter suas áreas acadêmicas qualificadas propriamente como das humanidades digitais). Assim, defendemos que os autores da área de *Information Studies* tem maior relevância no papel delineador do campo por receberem o maior número de citações de todos os agrupamentos de coautores, inclusive maior do que aqueles que mais publicaram.

As métricas de cocitação externam a proximidade entre autores, cuja proximidade é responsável por uma tessitura de uma área, de um campo ou, ainda, de uma disciplina, foi observada em seus níveis de autores, referências e fontes. Para a cocitação entre autores observamos que os grupos acadêmicos ora distantes, seguem em aproximação, sugerindo um movimento de consolidação da área. Pela disposição dos dados inferimos que o domínio começa a perder a identidade pluridisciplinar e a construir uma identidade própria do "novo" campo. Assim, podemos verificar que autores de grupos evidentes em outras métricas como 'computação' e 'informação', passam a dividir relevância com autores clássicos de áreas mais filosóficas. Destacamos ainda a constatação de novos autores terem despontado no ranqueamento por cocitação, além de observar também a reincidência de nomes já destacados neste grupo seletivo, promovendo, inevitavelmente, uma associação de influência no domínio por suas áreas acadêmicas - cenário também já observado em métricas anteriores.

A métrica de cocitações focada nas referências, revelou detalhes importantes sobre o campo através de sua literatura, nos permitindo supor que, apesar de toda a característica pluridisciplinar das humanidades digitais e seu apelo tecnológico, o campo da informação

(*InformationStudies*) assume um protagonismo ímpar por ditar com efusividade a temática das bibliotecas como as principais problematizações no domínio das HDs. A análise dessa métrica expõe com ineditismo a predominância absoluta de uma área numa métrica, ou pelo menos, não deixando margem de dúvida sobre essa interpretação. Assim, portanto, podemos inferir que a área de bibliotecas e informação tem um papel central nos debates em humanidades digitais, contribuindo para delinear e desenvolver o campo a partir de suas influências.

A lista de periódicos mais cocitados revelou que a principal a revista cocitada é da área de literatura e linguística, tendo o maior índice de cocitação. Existe ainda um grupo de destacada relevância com fontes nominadas como área de HDs enquanto que a disciplina de computação tem presença não tão destacada apesar de presente; a área da informação tem forte presença, todavia sem muita expressividade. Contudo o fato dos periódicos servirem de balizadores conceituais contundentes nos auxiliou nessa tarefa por serem instrumentos de comunicação formal dos resultados da pesquisa original para a comunidade científica, servindo como arquivo das ideias e reflexões dos cientistas, dos resultados de suas pesquisas e observações sobre os fenômenos da natureza.

Compreendendo o terceiro núcleo dos quais fizeram parte os autores e periódicos, procedemos também na análise de pontuar quais as palavras-chave possibilitam uma notória revelação sobre o escopo do campo e sobre sua real abordagem. Como sugerido pelo estudo, a estrutura proposta para análise das palavras-chave nos concedeu a identificação de áreas/campos, práticas/métodos e objetos/alvos esboçando uma estrutura temático-conceitual. Nesse sentido a estrutura temático-conceitual apontada pelo total de descritores de assunto pode ser categorizada enquanto as áreas: *education, semantics, history, social sciences, cultural heritage, information science, linguistics e computer science*. Já para práticas de pesquisa pudemos identificar: *humanities research, natural language processing systems, data mining, data handling, artificial intelligence, learning systems, engineering education, text mining, virtual reality, digital researches, collaboration e digitization*. Enquanto que objetos das humanidades digitais estariam identificados: *digital libraries, metadata, semantics, visualization, information systems, information retrieval, linked data, computational linguistics, ontology, artificial intelligence, digital collections, internet, information technology, research infrastructure, information management, algorithms, digital researches*.

Com efeito, o que se revela pelos termos não é apenas a estrutura na qual se sustenta o novo campo das *digital humanities*, mas principalmente a possibilidade de identificar, além de áreas eloquentes, quais contribuições estas fornecem juntamente com as práticas aplicadas aos objetos de natureza digital. O argumento desse estudo foi considerar quais termos encontraram-se com recorrência nas diversas categorias de palavras-chave para nos possibilitar realizar algum juízo de valor. Somou-se a essa recorrência o papel dos autores, enquanto agentes da construção de uma tessitura intelectual do campo emergente, não obstante a representatividade simbólica e intelectual fornecida pelos periódicos difundem essa literatura, influenciando também seus contornos intelectuais. Dessa maneira, podemos afirmar que algumas áreas efetivam sua presença no campo como, por exemplo, História e Educação enquanto outras tem participação não apenas como laboratório, mas como contribuintes de suas temáticas particulares como, exemplificamos, a Computação e a Ciência da Informação. Essa constatação nos permitiu identificar que enquanto algumas

áreas estão presentes nesse domínio como consumidoras dos produtos gerados pela tecnologia, outras contribuem com reflexões de suas práticas teórico-conceituais. Assim vimos que, enquanto áreas tinham sua presença nas humanidades digitais por estarem projetando novas ferramentas, outras contribuía para o campo com o somatório de implementação e criação de novas ferramentas, além de outras vendo suas reflexões e objetos serem alvo direto dessas inovações. Temos por exemplo nas práticas de pesquisa observadas com recorrência nas humanidades digitais como '*natural language processing systems*' e '*data mining*', além de '*visualization*', temáticas próprias da Computação. Ao passo que os objetos de pesquisa de maior reincidência apontamos '*semantics*', '*metadata*' e '*digital libraries*', temas próprios das discussões do campo informacional como a Ciência da Informação.

## Referências

- Chen, G., & Xiao, L. (2016). Selecting publication keywords for domain analysis in bibliometrics: A comparison of three methods. *Journal of Informetrics*, 10(1), 212–223. <https://doi.org/10.1016/j.joi.2016.01.006>
- De Bellis, N. (2009). *Bibliometrics and citation analysis: From the Science citation index to cybermetrics*. Scarecrow Press.
- Garfield, E. (1983). Science Citation Index: a new dimension in indexing: this unique approach underlies versatile bibliographic systems for communicating and evaluating information. *Essays of an Information Scientist*, n. 7, p. 525-535. <https://doi.org/10.1126/science.144.3619.649>
- Leydesdorff, L. (1998). Theories of citation? *Scientometrics*, 43(1), 5–25. <https://doi.org/10.1007/BF02458391>
- Marques, L. E., & Pinheiro, M. M. K. (2013). A Cúpula Mundial sobre a sociedade da informação - CMSI: Foco nas políticas de informação. *Informação & Sociedade: Estudos*, 23(1). <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/15450>
- Price, D. J. S. (1986) *Little science, big science and beyond*. Columbia University Press: New York, 1986. <http://garfield.library.upenn.edu/essays/v10p072y1987.pdf>
- Varela, A. V. (2005) A explosão informacional e a mediação na construção do conhecimento. In: ENCONTRO NACIONAL DE PESQUISA EM CIÊNCIA DA INFORMAÇÃO (ENANCIB), 6., 2005, Florianópolis, SC. Anais... <https://brapci.inf.br/index.php/res/download/173396>
- Zhao, D., & Strotmann, A. (2015). *Analysis and visualization of citation networks*. Springer.