

**Renan Marinho de Castro**

Análise da literatura das humanidades digitais: uma proposta bibliométrica para descrição de seu escopo e congruência conceitual

**Tese de Doutorado**

**Abril de 2020**



UNIVERSIDADE FEDERAL DO RIO DE JANEIRO - UFRJ  
ESCOLA DE COMUNICAÇÃO - ECO  
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA - IBICT  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO - PPGCI

**RENAN MARINHO DE CASTRO**

Análise da literatura das humanidades digitais: uma proposta bibliométrica para  
descrição de seu escopo e congruência conceitual

Rio de Janeiro

2020

UNIVERSIDADE FEDERAL DO RIO DE JANEIRO - UFRJ  
ESCOLA DE COMUNICAÇÃO - ECO  
INSTITUTO BRASILEIRO DE INFORMAÇÃO EM CIÊNCIA E TECNOLOGIA - IBICT  
PROGRAMA DE PÓS-GRADUAÇÃO EM CIÊNCIA DA INFORMAÇÃO - PPGCI

**RENAN MARINHO DE CASTRO**

Análise da literatura das humanidades digitais: uma proposta bibliométrica para  
descrição de seu escopo e congruência conceitual

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro/Escola de Comunicação, como requisito parcial à obtenção do título de Doutor em Ciência da Informação.

Orientador: Prof. Dr. Ricardo Medeiros Pimenta

Rio de Janeiro

2020

## CIP - Catalogação na Publicação

CC355a Castro, Renan Marinho de  
Análise da literatura das humanidades digitais:  
uma proposta bibliométrica para descrição de seu  
escopo e congruência conceitual / Renan Marinho de  
Castro. -- Rio de Janeiro, 2020.  
200 f.

Orientador: Ricardo Medeiros Pimenta.  
Tese (doutorado) - Universidade Federal do Rio  
de Janeiro, Escola da Comunicação, Instituto  
Brasileiro de Informação em Ciência e Tecnologia,  
Programa de Pós-Graduação em Ciência da Informação,  
2020.

1. Humanidades Digitais . 2. Análise  
bibliométrica . 3. Análise de domínio. 4. Revisão de  
literatura. 5. VOSviewer. I. Medeiros Pimenta,  
Ricardo, orient. II. Título.

**RENAN MARINHO DE CASTRO**

**Análise da literatura das humanidades digitais: uma proposta bibliométrica para  
descrição de seu escopo e congruência conceitual**

Tese de Doutorado apresentada ao Programa de Pós-Graduação em Ciência da Informação, convênio entre o Instituto Brasileiro de Informação em Ciência e Tecnologia e a Universidade Federal do Rio de Janeiro/Escola de Comunicação, como requisito parcial à obtenção do título de Doutor em Ciência da

Aprovada em 30 de abril de 2020.

---

Prof. Dr. Ricardo Medeiros Pimenta (Orientador)  
PPGCI-IBICT/UFRJ-ECO

---

Prof. Dr. Fábio Castro Gouveia  
PPGCI-IBICT/UFRJ-ECO

---

Prof. Dr. Renato Rocha Souza  
FGV

---

Prof. Dra. Gimena Del Rio Riande  
IIBICRIT-SECRET/CONICET (Argentina)

---

Prof. Dr. Marcelo Fornazin  
PPGCI-IBICT/UFRJ-ECO

À minha mãe por todo seu carinho e apoio.

Aos meus filhos, presentes celestiais, que espero amar, cuidar e proteger com absoluto zelo e esmero paternais.

À memória do pai que prometi honrar buscando o compromisso de ser um homem pautado pela retidão nas mais diversas instâncias da vida.

## AGRADECIMENTOS

A Deus, força suprema e onisciente, que nos concede a vida e a capacidade de seguirmos em frente mesmo quando tentados a crer que não somos mais capazes.

À minha mãe Lia Castro que, além de sempre me confortar nos momentos de esmorecimento, me prove sua intercessão e amor maternal acreditando em mim mais do que eu mesmo.

À minha esposa Jaqueline Castro por toda compreensão e tolerância nas ausências, por sua torcida para que tudo sempre alcançasse o sucesso.

À minha pequena filha Manuela Castro que teve o desafio de chegar em nossas vidas conjuntamente com o início dessa empreitada, minhas escusas por todas as oportunidades quais não pude contemplar com integralidade.

Ao meu filho Nicolas Castro, que, por sua vez, chegou concomitantemente ao nascimento desta tese, terá sua inocência recompensada com mais tempo para exercício das funções paternas deste pai.

À minha irmã e meu sobrinho-afilhado Fátima e Mateus Castro, cujas questionamentos pelo término dessa "ausência" podiam ser justificados pela não compreensão da íntegra deste processo de pesquisa.

À direção do Centro de Pesquisa e Documentação de História Contemporânea do Brasil da Fundação Getúlio Vargas, pela oportunidade e estímulo para meu desenvolvimento acadêmico.

À minha amiga e coordenadora Daniele Amado por toda parceria e compreensão às limitações impostas pelo compartilhamento do tempo entre o profissional e o acadêmico.

Aos meus colegas de trabalho que contribuíram com este processo pelo menos com suas compreensões e solidariedade com o andamento desta pesquisa.

Ao meu orientador, Ricardo Pimenta, honroso parceiro acadêmico e ser humano altamente sensível às imposições da vida pessoal e profissional, por sua compreensão, gentileza e empatia.

Aos meus colegas da turma 2016 do doutorado IBICT que essa jornada os fez mais próximos e parceiros.

Aos funcionários e professores do Instituto Brasileiro de Ciência e Tecnologia pelo suporte acadêmico ao longo desses anos.

Aos ilustres professores doutores que compõem essa banca pela honra de suas disponibilidades a quem os retribuo com admiração ímpar.

Ama e faz o que quiseres. Se calares, calarás com amor; se gritares, gritarás com amor; se corrigires, corrigirás com amor; se perdoares, perdoarás com amor. Se tiveres o amor enraizado em ti, nenhuma coisa senão o amor serão os teus frutos. (Santo Agostinho)

A característica do digital é tornar a abstração materialmente operacional, trazer essa abstração para o concreto, sem que ele deixe de ser abstrato. (EVENS, 2012)



CASTRO, Renan Marinho de. **Análise da literatura das humanidades digitais**: uma proposta bibliométrica para descrição de seu escopo e congruência intelectual. 200 f. 2020. Tese (Doutorado em Ciência da Informação) - Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2020.

## RESUMO

Esse estudo se propõe construir e visualizar mapas bibliométricos oriundos de uma revisão de literatura das humanidades digitais. O objetivo principal dessa pesquisa foi elaborar um mapeamento do campo através da construção de grafos para análise de métricas da bibliometria. Foi implementada uma leitura desses dados a partir da utilização do software VOSviewer para construção de um diagnóstico do escopo das humanidades digitais. A análise se centrou nos dados que fornecem subsídios para uma descrição temático-conceitual e do domínio das humanidades digitais ao analisar palavras-chave, autores e periódicos. Busca uma noção sensível daquilo que pode revelar a influência da produção de literatura acadêmica na composição das humanidades digitais. Aplica uma análise baseada em três pilares dos estudos bibliométricos: autores, periódicos e palavras-chave dos artigos acadêmicos 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. Combina conceitos acerca das abordagens de ordem bibliométrica com análises baseadas em ferramentas de visualização, legitimando práticas concernentes às humanidades digitais enquanto ferramenta de conhecimento do próprio campo. 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 de outras disciplinas nas humanidades digitais e visualiza com clareza que termos se apresentam com maior recorrência. Lista os termos mais recorrentes na literatura desdobrando essa listagem em termos indicados pelos autores e aqueles indicados pelos índices das bases de dados. Coteja esses dados como forma de alcançar uma expressividade temático-conceitual de um campo tido como interdisciplinar e híbrido. Categoriza o produto desses dados como forma de concluir, além de uma descrição temático-conceitual orgânica, um diagnóstico do domínio das humanidades digitais.

Palavras-chave: Humanidades digitais. Ciência da Informação. Análise bibliométrica. Análise de domínio.

CASTRO, Renan Marinho de. **Análise da literatura das humanidades digitais**: uma proposta bibliométrica para descrição de seu escopo e congruência intelectual. 200 f. 2020. Tese (Doutorado em Ciência da Informação) - Universidade Federal do Rio de Janeiro, Instituto Brasileiro de Informação em Ciência e Tecnologia, Rio de Janeiro, 2020.

## ABSTRACT

This study aims to build and visualize bibliometric maps from a literature review on digital humanities. The main objective of this research was to elaborate a mapping of the field through the construction of graphs for analysis of bibliometrics metrics. A reading of these data was implemented using the VOSviewer software to build a diagnosis of the scope of digital humanities. The analysis focused on the data which provide subsidies for a thematic-conceptual description and the domain of digital humanities when analyzing keywords, authors and journals. It seeks a sensitive notion of what can reveal the influence of the production of academic literature on the composition 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 the characteristics printed in the literature and only through it available. It combines concepts about bibliometric approaches with analysis based on visualization tools, legitimizing practices concerning digital humanities as a tool for knowledge of the field itself. It pursues a categorization capable of not only revealing details of its scope, but also providing elements indicative of which themes influence the area. It indicates the influence of other disciplines in the digital humanities and visualizes clearly which terms appear with greater recurrence. It lists the most recurrent terms in the literature, breaking down this list into terms indicated by the authors and those indicated by the database indexes. It compares these data in order to achieve a thematic-conceptual expressiveness of a field considered interdisciplinary and hybrid. It categorizes the product of these data to conclude, in addition to an organic thematic-conceptual description, a diagnosis of the digital humanities domain.

Keywords: Digital humanities. Information Science. Bibliometric analysis. Domain analysis.

## LISTA DE ILUSTRAÇÕES

Figura 1 – Esboço do mapa conceitual da <i>Humanities Computing</i> proposto por McCarty em 2003.....	52
Figura 2 – Fases na abordagem filosófica dos objetos digitais proposta por Hui (2002).....	58
Figura 3 – Esquema proposto enquanto prévia à concepção das humanidades digitais.....	65
Figura 4 – Proposta lógico-conceitual para as humanidades digitais.....	35
Figura 5 – Diagrama proposto para representação material das <i>Digital Humanities</i> .....	57
Figura 6 – Zonas de Distribuição da Curva de Zipf de acordo com Quoniam (1992).....	82
Figura 7 – Representação gráfica para formulação da Lei de Bradford.....	85
Figura 8 – Grafo da rede autores citados.....	117
Figura 9 – Grafo periódicos mais citados.....	121
Figura 10 – Grafo de documentos com maior número de links entre os mais citados.....	123
Figura 11 – Grafo de coautoria de autores.....	127
Figura 12 – Grafo de cocitação de autores.....	132
Figura 13 – Grafo de densidade de cocitação de autores.....	134
Figura 14 – Grafo de cocitação de referências - cluster 2 e 3.....	135
Figura 15 – Grafo de cocitação de fontes.....	138
Figura 16 – Grafo do acoplamento bibliográfico dos autores.....	143
Figura 17 – Grafo do acoplamento bibliográfico de artigos.....	146
Figura 18 – Grafo do acoplamento bibliográfico de periódicos.....	149
Figura 19 – Grafo de ocorrência de palavras-chave totais.....	154
Figura 20 – Grafo de ocorrência de palavras-chave de autores.....	159
Figura 21 - Grafo de ocorrência de palavras-chave do índice.....	163
Figura 22 - Grafo de instituições produtoras da literatura acadêmica.....	168
Figura 23 - Grafo de citação de países.....	172
Figura 24 - Grafo de citação de todos os países.....	174

Figura 25 - Grafo de coautoría países.....	175
Figura 26 - Grafo acoplamiento bibliográfico de países.....	179

## LISTA DE TABELAS

Tabela 1 – Sumário de possibilidades de análise para recursos bibliográficos.....	37
Tabela 2 - Autores mais citados por força de link.....	118
Tabela 3 – Periódicos mais citados por força de links.....	121
Tabela 4 – Documentos mais citados por força de links.....	124
Tabela 5 – Todos os autores em coautoria por força de links.....	128
Tabela 6 – Autores mais cocitados por número de citações.....	133
Tabela 7 – Referências cocitadas por força de links.....	136
Tabela 8 – Principais fontes cocitadas.....	139
Tabela 9 – Acoplamento bibliográfico por autor.....	144
Tabela 10 – Acoplamento bibliográfico por artigo.....	147
Tabela 11 – Palavras-chave 'totais' mais recorrentes (49 principais).....	156
Tabela 12 – Palavras-chave 'de autor' mais recorrentes.....	160
Tabela 13 – Palavras-chave 'do índice' mais recorrentes.....	164
Tabela 14 – Instituições ranqueadas por citação.....	168
Tabela 15 – Países ranqueados por citação.....	173
Tabela 16 – Países inversamente ranqueados por coautoria.....	177

## LISTA DE SIGLAS E ABREVIATURAS

ABA	Acoplamento bibliográfico de autores
CI	Ciência da Informação
DH	Digital Humanities
GIS	Geographic information system
HC	Humanities Computing
HD	Humanidades Digitais
RDF	Resource Description Framework
SCI	Science Citation Index
TaDiRAH	Taxonomy of Digital Research Activities in the Humanities
TEI	Text encoding initiative
TICs	Tecnologias da Informação e Comunicação

## SUMÁRIO

1 INTRODUÇÃO .....	19
1.1 Objeto.....	20
1.2 Justificativa .....	23
1.3 Marco teórico-conceitual .....	31
1.4 Objetivo.....	37
1.4.1 Objetivos Específicos.....	37
1.5 Metodologia .....	37
2 HUMANIDADES DIGITAIS COMO EXPOENTE DE UM NOVO CONCEITO	42
2.1 Prelúdio à conceituação das Humanidades Digitais .....	44
2.2 Objetos digitais e o não-finito: da maleabilidade à uma ontologia.....	48
2.3 As Humanidades Digitais além de um esboço conceitual .....	53
2.4 Arcabouço prático das Humanidades Digitais: uma parte da investida para compreensão do conceito .....	61
2.5 As Humanidades Digitais sob evolução epistêmica de sua própria heurística .....	66
3 UM FOCO NA ESTRUTURA ELITISTA DA CIÊNCIA PARA COMPREENSÃO DESTE DOMÍNIO: UMA RESSONÂNCIA DOS ESTUDOS BIBLIOMÉTRICOS	69
3.1 Uma abordagem bibliométrica como caminho para um diagnóstico das Humanidades Digitais.....	70
3.2 Ensaio de análise do domínio como produto do diagnóstico bibliométrico .....	79
3.3 A centralidade dos indicadores de mensuração do conhecimento: a formação da estrutura elitista da ciência.....	86
4 DESVELANDO AS HUMANIDADES DIGITAIS A PARTIR DA LITERATURA: UM ENSAIO PARA COMPREENSÃO DO DOMÍNIO EM RESSONÂNCIA À ABORDAGEM BIBLIOMÉTRICA.....	90
4.1 Constituindo uma matriz de dados para análise: recuperando a literatura sobre Humanidades Digitais.....	93
4.1.1 Elaboração da expressão de busca .....	99
4.1.2 Extraindo os dados recuperados pela busca: compondo a matriz de análise	101
4.2 Analisando as métricas e visualizando um campo: aplicando uma estratégia para compreensão do domínio das Humanidades Digitais .....	103
4.2.1 Análise de citações para o campo das Humanidades Digitais .....	105
4.2.2 Análise das fontes de publicação para o campo das Humanidades Digitais	110

4.2.3 Análise dos documentos mais citados na literatura das Humanidades Digitais	113
4.3 Análise de coautoria entre autores .....	116
4.4 Estruturando as cocitações para contribuição da percepção do campo das humanidades digitais.....	120
4.4.1 A cocitação entre autores .....	122
4.4.2 Referências de destaque nas cocitações: publicações mais influentes nas humanidades digitais .....	125
4.4.3 Complementando a relevância das referências cocitadas: analisando a cocitação de fontes .....	128
4.5 O papel do acoplamento bibliográfico para o domínio.....	132
4.5.1 Acoplamento bibliográfico de autores .....	133
4.5.2 Documentos referenciais na análise de acoplamento bibliográfico .....	136
4.5.3 Periódicos em acoplamento bibliográfico .....	139
4.6 Compreendendo a estrutura temático-conceitual do domínio: analisando as palavras-chave .....	142
4.6.1 Uma visão geral das palavras-chave .....	144
4.6.2 Palavras-chave dos autores.....	149
4.6.3 Palavras-chave da base de dados.....	153
4.7 Instituições produtoras da literatura.....	158
4.8 Abrangência da literatura .....	161
CONSIDERAÇÕES FINAIS .....	173
REFERÊNCIAS .....	187



## 1 INTRODUÇÃO

A informação tem assumido importância crescente na sociedade contemporânea, reconhecidamente associada ao poder econômico e cultural, o que ajuda a entender, em linhas gerais, como a entidade informação adquiriu o *status* que conhecemos atualmente. Sua projeção como uma engrenagem fundamental tanto quanto o próprio capital, criou as bases para o que hoje podemos denominar como sociedade da informação. Na visão de Capurro (2007) a sociedade atual atinge tal patamar, basicamente, ao estar caracterizada pela convergência desses fatores com as tecnologias da informação e seus impactos globais.

Alinhando-se a essa tendência, o movimento em curso levou ao desenvolvimento de abordagens pormenorizadas e, conseqüentemente, a estudos científicos de todo esse processo que se centra no objeto informacional. Em linhas gerais, temos alicerçadas aí as bases para o surgimento e aprimoramento da Ciência da Informação. Esse movimento deu corpo à ciência que teria a informação como seu principal objeto empírico, e a área que se delimitaria desde os anos 60, por fundamento, no conceito do termo ‘informação’ (SOUZA, 2013). Mesmo ainda estando passando por um natural processo de consolidação conceitual, a área projeta-se também por estar intimamente ligada a uma realidade interdisciplinar inspirada em técnicas e práticas consolidadas de áreas correlatas. Dessa forma, a Ciência da Informação debruça-se sobre o estudo da informação de sua gênese até sua consolidação em forma de conhecimento, passando assim por todas as etapas de um processo complexo como análise, coleta, classificação, manipulação, armazenamento, recuperação e disseminação da informação.

Apesar da área ser relativamente recente, a Ciência da Informação vem se consolidando fortemente com a reflexão de pensadores tradicionais e contemporâneos, mas principalmente por uma atuação empírica no campo informacional. Esse movimento deu corpo à abordagem matemática do estudo da informação, remodelou conceitos e marcou definitivamente a área informacional com uma ótica peculiar, própria de uma Ciência nascida para dar conta desse fenômeno, a despeito de outras formas de trato informacional como a computação ou a comunicação.

Os problemas básicos de se compreender a informação e a comunicação, suas manifestações, o comportamento informativo humano e os problemas aplicados ligados ao "tornar mais acessível um acervo crescente de conhecimento", incluindo as tentativas de ajustes tecnológicos, não podem ser resolvidos no âmbito de uma única disciplina. Este fato ficou claro, a partir da afirmação de BUSH, para todos que refletiram acerca das complexidades

envolvidas. Problemas complexos demandam enfoques interdisciplinares e soluções multidisciplinares. (SARACEVIC, 1996, p. 48)

Não se portando indiferente ao seu gene interdisciplinar<sup>1</sup>, a Ciência da Informação vem se desenvolvendo atenta a esta vocação e inspirando-se por essa natureza que o autor categoriza como não esgotável no âmbito de uma única disciplina. Amparados por essa função, utilizamo-la como razão pela qual se pretendeu contextualizar nesta pesquisa sua relação interdisciplinar com áreas que possuem a mesma inclinação, sobretudo diante de uma massiva penetração tecnológica que transformou completamente o cenário de convivência científica. Transformação esta que para Castells (2011) tem o ponto de partida em uma mudança de paradigma situada no final do século XX, que desde então se pontua um novo momento da história cuja característica é um modelo tecnológico que se organizou em torno da tecnologia da informação. Esse processo compreende uma tentativa de acompanhar de perto a evolução de ferramentas que possibilitem o melhor aproveitamento dos produtos das TICs. Assim, inevitavelmente, as transformações da sociedade da informação se fizeram sentir, de modo especial, na Ciência da Informação, já que o trato das informações, característica crucial dessa área, sofreu fortes transformações em termos de quantidade, suporte e infra-estrutura. A área obrigou-se a uma adaptação às mudanças radicais como às novas ferramentas informáticas e a virtualização de suportes.

## 1.1 Objeto

Nesse contexto de produção exponencial de informação, seguido de perto pela profusão de ferramentas computacionais aplicadas às metodologias de pesquisa num ambiente alheio à funções dessa ordem, se acha espaço para uma nova reflexão. Uma reflexão oriunda de indivíduos que experimentaram a metodologia tradicional elevada a uma nova dimensão, na qual a tecnologia é a maior responsável pela execução da função árdua, cuja comunidade envolvida vem denominando de *digital humanities*. Assim, por considerarmos que a expressão mais genuína da nova dimensão metodológica surge como

---

<sup>1</sup> Para Saracevic (1996, p.48) a interdisciplinaridade foi introduzida na CI pela própria variedade da formação de todas as pessoas que se ocuparam com os problemas descritos. Entre os pioneiros havia engenheiros, bibliotecários, químicos, linguistas, filósofos, psicólogos, matemáticos, cientistas da computação, homens de negócios e outros vindos de diferentes profissões ou ciências. Certamente, nem todas as disciplinas presentes na formação dessas pessoas tiveram uma contribuição igualmente relevante, mas essa multiplicidade foi responsável pela introdução e permanência do objetivo interdisciplinar na CI.

campo de conhecimento, delineado pelas experimentações praxiológicas sob o rótulo de humanidades digitais, cremos haver frutuosa relação 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. Assim, tal característica fundamental nos serve de mote para buscarmos tomar como objetivo desta pesquisa uma relação de similitudes potencialmente contributivas com o que vem atendendo por *digital humanities*. Visando, assim, uma análise perceptiva da congruência de escopos a partir da interpretação do domínio das humanidades digitais se inspirando, ainda, numa mensuração da influência de outros campos acadêmicos nas *digital humanities*.

As humanidades digitais (HDs) vêm se projetando no mesmo contexto da sociedade da informação diante da realidade da explosão informacional com foco nas fontes de informação digitais. A área recém-criada, se apresenta como um campo importante de interlocução que se desenvolve rapidamente através da crescente criação de centros acadêmicos para sua reflexão e implementação de projetos desta ordem. Tem seu marco de fundação pela promulgação do 'Manifesto das humanidades digitais', instrumento que cunha oficialmente a nomenclatura. Esse documento esclarece a motivação transdisciplinar da área que está fundamentada nas novas formas de uso dos registros digitais, ou como contextualiza o Manifesto (2010) "na opção da sociedade pelo digital que altera e questiona as condições de produção e divulgação dos conhecimentos". O documento também constata "que o digital induz uma presença mais forte dos aspectos técnicos e econômicos na pesquisa; que esta obrigação é uma oportunidade para fazer evoluir o trabalho coletivo". Em linhas gerais a área se dedica às novas utilizações das fontes de informação no formato digital e à reflexão sobre esse movimento, de modo especial, nas áreas temáticas próprias de sua atuação. Numa abordagem geral acerca das HDs podemos considerar que

As humanidades digitais referem-se ao conjunto das Ciências Humanas e Sociais, às Artes e às Letras. As humanidades digitais não negam o passado; apoiam-se, pelo contrário, no conjunto dos paradigmas, no saber fazer e conhecimentos próprios dessas disciplinas, mobilizando simultaneamente os instrumentos e as perspectivas singulares do mundo digital. As humanidades digitais designam uma transdisciplina, portadora dos métodos, dos dispositivos e das perspectivas heurísticas ligadas ao digital no domínio das Ciências Humanas e Sociais. (MANIFESTO das Humanidades Digitais, 2010, não paginado.)

Podemos considerar que as HDs surgem para categorizar uma realidade contemporânea de presença tecnológica no âmbito das fontes tradicionais de informação,

antes usufruídas apenas em seu formato físico. Trata-se de um movimento que, ao impactar a área das ciências humanas e sociais, levam os pares a percepção de que as pesquisas agora passam a ser mediadas pelas tecnologias. Amparadas pela digitalização, a tendência irreversível de criação de fontes digitais colocou às ciências humanas o desafio de incorporar novos métodos à sua tradicional metodologia de pesquisa. Esse desafio implica em uma nova forma de trabalhar, tornando essencial sua reflexão pelos envolvidos nela.

Essa aproximação pode ser sentida e atribuída, inicialmente, a uma lógica de interposição de objetos de estudo baseada no advento do digital. O elemento digital oferece à humanidade um novo paradigma, elevando com isso a uma confluência de ações, abordagens e interesses. Esse fenômeno nos serve de ponto de partida de um compartilhamento de reflexões plurais, levando diversos agentes a um ponto de encontro. Dessa forma o cenário contemporâneo da sociedade da informação, aliada à digitalização desenfreada, conferem aos objetos digitais uma posição crucial nesse cenário criando novos debates em torno da preservação digital, gestão da informação no âmbito da computação e no que tange a interposição de objetos de estudo. Com importância crescente e sua gravitação em áreas diversas, é crucial imputar à Ciência da Informação demarcar uma abrangência de sua atuação nesse fenômeno, atribuindo ao conceito e seu escopo um indicativo dessa interoperabilidade.

Por sua natureza, o termo objeto digital pode ser compreendido como equivalente aos suportes tradicionais da informação, uma vez que estes têm sido senão substituídos, elevados em importância por sua maleabilidade, corroborando assim nossa ideia de um objeto compartilhável. Não nos surpreende a constatação que a presença de arquivos de computador (*file*), documentos digitais, artefatos digitais, recursos digitais e materiais digitais vêm superando em escala exponencial os suportes tradicionais. Para Yamaoka; Gauthier (2013) as diferentes nomenclaturas são convergentes no sentido de referir-se a entidade 'informação' em sua ausência de fisicalidade. Assim os autores consideram como definição geral que objeto digital é um objeto, sobretudo, de informação, de qualquer tipo e formato expresso sob a forma digital nos servindo, portanto, como um núcleo ambíguo.

Os objetos digitais são marcados por um conjunto limitado de atributos que lhes conferem um perfil distinto e funcional. Os atributos dos objetos digitais que os diferem de objetos físicos são quatro: 1) a editabilidade que é uma característica intrínseca dos objetos digitais e pode ser alcançada suprimindo, adicionando, modificando elementos ou fazendo a atualização regular e contínua de itens ou campos; 2) a interatividade que é sua principal qualidade,

através da qual o agente humano pode ativar funções incorporadas no objeto; 3) aberto e reprogramável pela possibilidade de ser modificado por outro objeto digital; e 4) distribuídos por raramente estar limitado a uma única fonte, portanto, as fronteiras originalmente não existentes são criadas e mantidas tecnologicamente (KALLINIKOS; AALTONEN; MARTON, 2010 *apud* YAMAOKA; GAUTHIER, 2013, p.85-86).

Por sua complexidade os autores consideram que o objeto digital possui três propriedades herdadas das classes física, lógica e conceitual. Em sua fisicalidade defendem que são uma inscrição de sinais em uma mídia, ou seja, constituem-se em dados gravados num suporte físico que pode estar localizado localmente ou em outro espaço geográfico. Como lógico entende-se sua capacidade de mesmo descodificado e recodificado em relação à linguagem humana, está perfeitamente compreensível a um *software* que reconhecerá os *bits* e *bytes* neles impressos. Em sua classe conceitual, equipara-se à informação tradicional, permitindo ser lido e compreendido pela ação humana, configurando-se como dado passível da geração de conhecimento tal qual um documento tradicional. Assim "é um objeto 'do mundo real', reconhecido como uma unidade significativa de informação, tal como um livro, um contrato, um mapa ou uma fotografia" (YAMAOKA; GAUTHIER 2013, p.83).

Dessa forma evidencia-nos uma inclinação na noção geral sobre as humanidades digitais, amparadas no surgimento do digital, vertentes implícitas na direção dos estudos informacionais. Essa ideia, somada à interdisciplinaridade, bem como problemáticas outras que insurgem reflexões críticas nas HDs que possuem respaldo na *expertise* da Ciência da Informação, nos motiva à busca de aproximações entre as duas disciplinas. Ainda assim sua existência suportada na questão do digital remete as áreas a um núcleo comum, a uma interseção que sobrepõe as duas lógicas conceituais. Nesse sentido a pesquisa se inspira na percepção de uma visão sobre os estudos dos usos da informação na nova realidade tecnológica e digital que parecem refletir, pelas humanidades digitais, em uma abordagem dessa temática realizada externamente à Ciência da Informação. Assim nossa proposta pretendeu criar bases para uma visão perceptiva e interpretativa da confluência de escopos deste domínio em busca de algum tipo de identificação do alcance dos estudos informacionais no núcleo das humanidades digitais.

## 1.2 Justificativa

A hipótese de uma congruência nos domínios e escopos de uma área "nova" baseada em reabordagens de problemáticas atuantes em campos já consolidados, sugere aproximação entre temática e práticas das humanidades digitais e ciência da informação. Com os ensaios reflexivos em humanidades digitais apontando, com clareza, duas vertentes: uma estritamente reflexiva e outra de ordem praxiológica, que neste caso tangem, logicamente, os projetos calcados em utilização de ferramentas, devem haver, também, contribuições de ordem operacional oriundas de contribuições das ciências da informação. Além das reflexões sobre o digital, essas possíveis agregações se apresentam no âmbito do estudo do comportamento informacional que, no caso das humanidades digitais, devem ir ao encontro da problematização acerca do uso do ferramental computacional como mediador da informação. Mesmo uso esse que se credencia enquanto cerne, tantas vezes exposto, do fundamento da área das *digital humanities* merece ser colocado sob problematização. Essas reflexões têm amplo abrigo em pesquisas de ordem das competências em informação, por exemplo, sobretudo seus desdobramentos cuja sintonia fina encontra eco nas competências digitais que constituem marcas indeléveis da Ciências da Informação.

Com isto posto, podemos tomar por base o manifesto, no qual constatamos a principal definição desta área que reconhece a interdisciplinaridade das HDs como característica fulcral do campo, tanto para sua constituição quanto ao seu desenvolvimento. Parte desse alicerce a construção da ideia de que indivíduos equipararam suas perplexidades pela nova possibilidade de trabalhar fazendo multiplicarem experimentações no domínio do digital nas ciências humanas nos últimos anos. Fazem assim emergir, espaços acadêmicos dedicados à produção desses novos produtos, estabelecendo-se como protótipos ou lugares de aplicação de uma abordagem das *digital humanities*. A pluralidade desses indivíduos viu, portanto, na emergência das humanidades digitais, o espaço apropriado para uma reflexão uníssona desse movimento, pois

Existem múltiplas comunidades específicas, oriundas de interesses por diversas práticas, instrumentos ou objetos transversais (codificação de fontes textuais; sistemas de informação geográfica; lexicometria; digitalização do patrimônio cultural, científico e técnico; cartografia da web; mineração de dados; 3D; arquivos orais; artes e literaturas digitais e hipermediáticas; etc.) e que estão convergindo atualmente para formar o campo das humanidades digitais. (MANIFESTO das Humanidades Digitais, 2010, não paginado)

Para Russel (2011), por exemplo, as HDs são um termo para esse novo campo interdisciplinar que procura compreender o impacto e a relação de tecnologias da computação no trabalho de pesquisadores na área de humanas. Assim como um rótulo, uma classificação que intenciona caracterizar um movimento ou processo que até então não dispunha de uma categorização que a pudesse identificar com relativo ineditismo. Dessa forma, nos facilita perceber que a área surge a partir da efetivação do uso das tecnologias, de modo especial, àquelas para proficuidade de registros digitalizados e o ferramental que trouxe em seu bojo um novo cenário de pesquisa para os pesquisadores das ciências humanas. Mas não apenas reduzida a isso, as humanidades digitais podem ser consideradas,

um campo de estudo, pesquisa, ensino e inovação preocupados com a interseção da computação com as disciplinas da área de humanas. É por natureza metodologicamente interdisciplinar em sua abrangência. Tratando de pesquisa, análise, síntese e disponibilização da informação em formato eletrônico. Estuda como estes elementos afetam grande parte das disciplinas em que se encontram presentes e o que estas disciplinas têm a contribuir para o nosso conhecimento em formato computacional. (KIRSCHENBAUM, 2010, p.2, tradução nossa).

Mesmo caminhando a passos largos em termos de produção acadêmica focada em teorizar esse novo campo, as humanidades digitais ainda concentram-se nas práticas de implementação de projetos de disseminação de conteúdos digitais. Essas práticas estando ou não inseridas em contextos informacionais como as bibliotecas ou centros de documentação, podem usufruir de uma abordagem característica da área dos estudos da informação. Desta maneira vemos que o principal nicho das *digital humanities* e grande parte de seu espaço de nascimento constituem terreno pertencente às ciências da informação. Essa comunhão de interesses de análise é selada no âmbito do objeto digital que marca os pontos cardeais da pesquisa sob essa nova realidade. Assim, assumimos, portanto, que grande parte desses projetos de *DHs*, sobretudo àqueles abrigados em unidades de informação, trazem em seu "dna" características da Ciência da Informação.

Autores como Fitzpatrick (2012) defende que as humanidades digitais caracterizam-se por uma robusta área de pesquisa e ensino, que também está centrada não apenas na prática do digital, mas igualmente no compromisso de teorizar esse movimento. Para ele a área estaria marcada por um distanciamento entre os que se envolvem com a prática da área e, por outro lado, aqueles que se dedicam à interpretação desses novos movimentos. Segundo ele ainda existe uma tensão entre os que identificam a *digital humanities* como uma divisão entre estudiosos que utilizam as tecnologias digitais no

estudo de objetos tradicionais e aqueles que usam os métodos das HDs apenas no estudo de objetos digitais. Essa tensão explícita, a nosso ver, além da necessidade de um debate sobre o amadurecimento do campo, uma discussão reflexiva sobre questões de capacitação para uso de ferramentas, em sua predominância computacional, cada vez mais complexas. Exigindo assim conhecimentos alheios aos campos e convidando as HDs a contemplar em seu escopo o debate sobre, por exemplo, competências em informação e estudos de perfil de usuários, temáticas intimamente ligadas às ciências da informação.

Apesar da abundante vocação interdisciplinar, o compartilhamento de um mesmo objeto de estudo e as reflexões próprias da área adequadas às HDs há pouca relação com vistas a uma aproximação problematizada entre a Ciência da Informação e as humanidades digitais. O que nos motivou compreender a justaposição entre essas áreas do conhecimento, para então, aperfeiçoar o entedimento sobre esse relacionamento.

Com a disseminação desenfreada de conteúdos digitais, são colocados à prova os métodos de mensuração da produção do conhecimento, bem como o objetivo de atingir o indivíduo. Para Gómez (2009) a documentação digital desafia o cenário tradicional da disseminação da informação, porque para ela, a autonomia dos sujeitos passa a carecer de intencionalidade no processo de comunicação ou troca de “informações”, "característica essencial do cenário tradicional, está ausente na geração de um vasto conjunto de enunciados digitais já que sua produção e processamento ocorrem fora da consciência." (FROHMANN *apud* GÓMEZ, 2009 p. 125-126.)

Com a informação desgarrada de seus repositórios históricos tradicionais como bibliotecas e arquivos, o digital inevitavelmente confere uma descontextualização severa da informação que antes exigia uma proximidade, por exemplo, com lugares de memória. Esse fator se exacerba se considerarmos que uma área surja para configurar esse processo tal como ele está, ou seja, corroborando o uso da informação apenas em seu gene digital, distanciada dos processos informacionais complexos estudados pela Ciência da Informação que se dedicam desde a gênese do suporte.

Na medida em que o digital vira atributo de uma pluralidade de modos do fazer e do viver, e de seus produtos e resultados, vários conceitos buscarão descrever e dar uma denominação generalizante a esses processos de contornos e extensão incalculáveis. Trata-se de macro-conceitos. Ao mesmo tempo em que são descritivos, são fornecedores de um domínio de legitimidade à expansão das TIC's e seus efeitos de digitalização. **Será demandado um novo horizonte ontológico para as questões de informação**, a qual não poderia já ser igualada a um objeto discreto (coisa ou artefato), nem meramente encapsulada num subsistema funcional das organizações (um serviço, um agregado de informação). (GÓMEZ, 2009 p. 126, grifo nosso).



A autora, ao mesmo tempo em que vislumbra o surgimento de novas denominações para esse fenômeno, faz imperativa a abordagem de novos conceitos na própria Ciência da Informação. As humanidades digitais podem ser consideradas como expressão desse surgimento, porém formalmente desvinculado da Ciência da Informação, enquanto a nosso ver, por outro lado, o "novo horizonte ontológico" para este fenômeno urge ser moldado na perspectiva da área apesar de ainda não estar mapeado nem identificado.

Consideramos, ainda, que um "vício" pela tecnologia foi o responsável por alterar o patamar da lida com a informação pelo campo informacional. Nosso argumento se sustenta sob a interligação dessas tratativas sociais convergidas no que classificamos de objetos digitais. Os objetos digitais são os responsáveis por congregarem diversos campos e indivíduos sob um mesmo prisma, passando o foco de suas atividades profissionais e sociais para o dado binário expressão computacional do digital que pode ser armazenado, preservado, transmitido e utilizado de formas mais eficazes que outrora.

Essa digitalização<sup>2</sup> desenfreada aglutinou o interesse de várias áreas de estudos aproximando-as de um mesmo núcleo duro. Atualmente todos os envolvidos com práticas de informação, memória e cultura devem ter preocupações semelhantes como, por exemplo, a grande questão do armazenamento e o acesso dos dados digitais mediados, obrigatoriamente, pela tecnologia. Seja disponível em sites ou repositórios, as preocupações estão muito próximas das reflexões sobre as competências individuais.

Essa proximidade entre áreas, antes distante, começa a esboçar-se conceitualmente no que se vem classificando como humanidades digitais. As humanidades encontram no digital o elo que os vinculam em torno das mesmas questões. Há agora um comprometimento em como fazer uso de todos esses novos recursos, como a nuvem ou dados associados a posicionamento geográfico, fortalecendo assim esse argumento de que o conhecimento, ora manejado pelas áreas em particular, agora passa a ser compartilhado num mesmo canal sob a forma de dados de codificação binária, representados para nossa compreensão em versões digitais acessíveis somente a partir de um dispositivo capaz de acessá-lo, interpretá-lo e exibi-lo. Para que tamanha tendência irreversível não se configure num fator impeditivo, as contribuições das reflexões desenvolvidas no âmbito das competências informacionais são igualmente urgentes para

---

<sup>2</sup> Considerado aqui tão somente enquanto processo através do qual um dado analógico é tornado digital.

que esse processo possa significar um avanço no sentido da ampliação na disseminação da informação e do conhecimento.

O ambiente informacional passou por fortes transformações que não foram acompanhadas por todos os indivíduos no mesmo ritmo. As formas de leitura de um jornal ou um livro, por exemplo, não são as mesmas de anos atrás o que atribuiu aos dispositivos uma importância crucial na mediação, sobretudo nos ambientes informacionais. Se as transformações pelo uso da tecnologia impactaram toda a atividade humana, no âmbito informacional não foi diferente. Nesse contexto surgem as reflexões em torno da competência em informação que dá conta desse processo em sua amplitude, não deixando de receber atenção e importância ímpar em toda a área da Ciência da Informação. Segundo Miranda (2006) a noção da aplicação dessa percepção

Podem ser caracterizadas duas correntes tratando do assunto: 1) uma referente à gestão de recursos humanos, que preconiza o uso da competência para integrar as atividades de gestão de recursos humanos por meio do seu uso em processos de seleção, treinamento, avaliação e remuneração; 2) outra referente à sociologia da educação e do trabalho, que trata dos aspectos psicossociais da utilização da competência nos programas educacionais nos níveis de qualificação da mão-de-obra e de emprego (Guimarães, 2000 *apud* MIRANDA, 2006, p. 106).

Para a autora o processo em torno da competência em informação foi bastante amplo e generalista até que chegasse às abordagens obtidas no campo da Ciência da Informação. Até então as competências eram entendidas inicialmente enquanto capacidades técnicas de ordem organizacional voltadas quase que exclusivamente para o trabalho. Uma vez que "autores norte-americanos tratam a competência como um estoque de qualificações que credencia o exercício de determinado trabalho e os autores franceses associam a competência às realizações da pessoa em determinado contexto". (MIRANDA, 2006, p.107). Apenas posteriormente as reflexões em torno das competências passaram a se dar no campo da intelectualidade. A autora também entende que a migração para uma abordagem cognitiva foi conduzida de maneiras próprias, porém convergentes entre pensadores norte-americanos e franceses, assim segundo ela essa construção se conduziu da seguinte forma

Na abordagem americana, McClelland (1973) começou a estruturar o conceito de competência na década de 1970 questionando os testes de aptidão intelectual. O autor observou que as medidas de proficiência no trabalho dependiam de vários outros fatores, além da inteligência. A competência seria uma abordagem alternativa de avaliação da performance, em vez dos testes de inteligência. Boyatzis (2004) questionou outra ideia, também corrente, de que uma gerência efetiva estava diretamente relacionada à quantidade adquirida de conhecimento. As competências seriam características que proporcionam resultados efetivos – incluindo habilidades cognitivas ou intelectuais –, habilidades intrapessoais e habilidades interpessoais. Além do conhecimento e

das competências, uma performance acima da média requer o desejo de desenvolver e usar os talentos; e esse desejo é dirigido pelos valores, crenças, motivos e características pessoais. Na abordagem francesa, **ter competência é saber mobilizar e combinar recursos (pessoais e do meio). É, também, a faculdade de usar essa dupla instrumentalização de recursos de maneira pertinente, ou a capacidade de integrar saberes diversos e heterogêneos para finalizá-los na realização de atividades. A lógica de integração do saber, do saber-fazer e dos comportamentos se estabelece em função das exigências da situação de trabalho.** (MIRANDA, 2006, p. 107) Grifo nosso

A nosso ver essa abordagem foi determinante para apropriação do conceito pela área informacional, haja vista que a grande dependência ferramental está intimamente atrelada à um apelo cognitivo nas demandas por informação. A autora expõe essa visão por nós defendida ao aproximar as reflexões sobre competência em informação do conceito de necessidade informacional. Como sabemos a necessidade de informação é a mola propulsora dos estudos do trato informacional por garantir que toda a informação seja tratada segundo parâmetros de disponibilidade para recuperação por indivíduos que apresentam alguma motivação em obtê-la. Assim, julgamos que, atualmente, se faz impositivo a agregação da heurística dessas abordagens voltadas a uma aplicação nas *digital humanities*.

A autora atrela às largas reflexões sobre estudos de usuários a um núcleo mais refinado no tocante ao conceito. Para ela há um refinamento dessa temática no sentido dos estudos das necessidades de informação dos usuários. A autora afirma ainda que, Dervin e Nilan (1986) concluíram ser necessário mudar o paradigma tradicional e desenvolver uma forma alternativa para os estudos de necessidades e usos da informação. Nesse movimento Miranda (2006) identifica que

Foram detectados novos direcionamentos apontando para as seguintes tendências: as necessidades dos usuários deveriam se tornar o foco central da operação de sistemas; os serviços de informação deveriam ser ajustados às necessidades específicas do indivíduo, e não o contrário; deveria ser mudado o foco dos sistemas de informação dirigidos a tecnologias e conteúdos para os dirigidos aos usuários; deveria ser colocado o foco nos próprios usuários. (MIRANDA, 2006, p.100).

Por essa abordagem podemos especular que a grande mudança de foco nas reflexões em competência em informação que estimularam sua aproximação com a CI se deu a partir do momento em que essas competências construíram suas bases no aspecto cognoscível dos indivíduos. Para a autora, nesse contexto, os estudos orientados para o usuário nas abordagens mais recentes veem "a informação como uma construção subjetiva, em que o valor da informação reside no relacionamento que o usuário constrói entre si mesmo e determinada informação" (MIRANDA, 2006, p. 112). Assim o antigo

estereótipo de usuários, bem como seus estudos, passou por uma atualização na qual se considera o caráter próprio de cada indivíduo e o que lhe é capaz de prestar auxílio por si só em sua busca da satisfação pelas suas necessidades informacionais. Esse é, a nosso ver, a mais substancial qualidade de associação dos estudos das competências para e na Ciência da Informação. Por essa razão julgamos se fazer imperativo conscientizar as reflexões em humanidades digitais dessa necessidade, haja vista sua característica primacial, que delega integralmente ao computacional a mediação entre informação e indivíduos. Nesse sentido, vislumbramos em tal problemática mais um argumento incisivo para estimular ainda mais a congruência de escopos dos dois campos.

Identificar uma maneira de experimentar a competência em informação que depende da disponibilidade e usabilidade das tecnologias de informação. Visto desta forma, a pessoa capacitada informacionalmente é aquela que examina o ambiente informacional para alcançar um alto nível de conhecimento de informação. De acordo com este ponto de vista, é possível experimentar a competência em informação se você for membro de uma comunidade que suporte o uso de tecnologias. (BRUCE, 2003, p. 290, tradução nossa).

Como se refere a autora, a tecnologia cria uma forma de agrupar pares capazes de interagir com tais instrumentos computacionais, que por sua vez, reúne os indivíduos em comunidades. Um exemplo genérico dessa ideia seria a interação em redes sociais, no qual apenas os indivíduos capazes de interagir com uma determinada plataforma social podem agrupar-se a ela e, por consequência, ter acesso aos dados ali compartilhados. Logo, os indivíduos que não detêm essa competência estão de fora desse convívio e do desfrute dos dados lá compartilhados. Portanto, se faz cada vez mais importante as reflexões em torno das competências em informação no contexto informacional e as contribuições nesse sentido, que igualmente serviriam à própria área das HDs, ao promover uma modalidade de auto identificação, contribuindo a uma pormenorização da fragmentação oriunda da pluralidade dos pares. Assim a problemática das competências informacionais também estariam aptas a qualificar uma aproximação dos campos que já a desenvolvem para uma contribuição no amadurecimento das humanidades digitais.

Por outro lado, para Tang; Cheng; Chen (2017) as HDs apresentam um terreno fértil e apropriado para o estudo da integração do conhecimento por focar diretamente nos problemas pragmáticos de como a computação auxilia na prática de pesquisa e como problemas teóricos e práticos recebem nova perspectiva através da computação. Segundo os autores,

A integração cognitiva de conceitos, teorias, métodos e/ou resultados de diversos campos são considerados a marca registrada da pesquisa interdisciplinar [...] A integração do conhecimento é um processo

caracterizado pela alta heterogeneidade (diversidade) cognitiva e pelo aperfeiçoamento da estrutura relacional (coerência), na qual a coerência é definida pela extensão de assuntos específicos, conceitos, ferramentas e interconexão. (TANG; CHENG; CHEN 2017, p.2, tradução nossa).

Nesse sentido, compactuamos com a tese dos autores de que diferentemente de outras áreas, as *digital humanities* vivenciam, por sua natureza, uma convergência de reflexões que devem conduzir o campo para uma formatação própria dentro de um processo de consolidação típico da evolução do conhecimento. Assim entendemos necessário agregar perspectivas informacionais à nova área levando para lá vieses e performances de uma área senão tão íntima, tematicamente próxima, como as ciências da informação.

Ao constatar que existem diversos projetos do campo que são tocados por indivíduos ligados à Ciência da Informação sem atender pelo nome humanidades digitais, se constituiria o mote para uma prévia, porém não identificada e conseqüentemente não mensurada, interlocução das áreas. Com efeito, não se pode desconsiderar as contribuições que merecem ser prestadas também pelos estudos de usuários, por exemplo, que enquanto objetivo final dos esforços de disponibilização de informação, que também situam-se à margem das problemáticas desenvolvidas no âmbito das HDs.

Somando-se a uma expectativa de identificação de um campo a partir de sua característica básica de integração do conhecimento, cientes de que existe uma relação de reciprocidade no que tange seus objetos, podemos, então, como pontuam Tang; Cheng; Chen (2017), prospectar reflexões no âmbito da análise de domínio. Propondo, assim, o uso de técnicas de pesquisa da Ciência da Informação para extração de conhecimento passível de contribuição a um diagnóstico do campo das humanidades digitais e sua não hipotética interlocução com a Ciência da Informação. Dessa forma uma percepção partindo de diagnóstico do domínio, amparada por um estudo bibliométrico, deve contribuir para a construção de conhecimento sobre esse campo. Amparando-se assim no gabarito da Ciência da Informação em reflexões em torno da interdisciplinaridade para mirar em algum tipo de checagem de fronteiras relacionais dessa nova área.

### **1.3 Marco teórico-conceitual**

Pela natureza do domínio aqui tido como objeto de estudo, onde seu *status quo* reflete uma lógica de consolidação do próprio campo, é sensível propor uma reflexão que

parta de estudos consolidados da relação entre a área das HDs e as tidas em interdisciplinaridade. Dessa forma a estratégia desta pesquisa foi aplicar conceitos próprios da área informacional à contribuição daquela sem desprestigiar a relação entre as duas. Assim a pesquisa assumiu balizadores conceituais característicos da Ciência da Informação, por exemplo, para aplicar seus modelos às *digitais humanities*. Propomos, portanto, desvelar relações estreitas tendo enquanto argumento de pano de fundo a sobreposição de objetos e a convergência de escopos. Dessa maneira a pesquisa visou utilizar de modelos bibliométricos enquanto instrumento adequado de análise da literatura, discutir nesse contexto contribuições da ordem de análise do domínio e por fim aproximar os debates sobre a interdisciplinaridade que contribuíram na consolidação deste campo, para agregar valor, senão às humanidades digitais, pelo menos à relação desta com a Ciência da Informação.

As áreas das humanidades possuem grande inclinação para o uso da recuperação de literatura científica produzida sobre um determinado tema, para, ao tentar identificar seu *status quo*, embasar pesquisas atuais. A Ciência da Informação faz desta prática um instrumento de condição *sine qua non* para preceder pesquisas em seu domínio, identificar estudos e pesquisadores relevantes, bem como conhecer o estado-da-arte de determinada temática. Fato é que pela praxe que este instrumento assumiu, sua importância o fez requintar sua metodologia no sentido da agregação de credibilidade a esse instrumento.

Para SEGURA-MUÑOZ; TAKAYANAGUI; SANTOS et al (2013, p.1) o instrumento conhecido por revisão de literatura "apresenta um caráter descritivo-discursivo, caracterizando-se pela ampla apresentação e discussão de temas de interesse científico" permitindo ao leitor adquirir e atualizar o conhecimento sobre uma temática específica de maneira concreta em um intervalo de tempo relativamente curto. Todavia, o instrumento consolidado por revisão de literatura, carecia de estratégias metodológicas que dificultavam sua re-edição, além de uma incapacidade de contemplar em seu alcance questões pontuais, devendo em expectativa quando a intenção da revisão partia de uma hipótese mais complexa de pesquisa.

Dessa forma vimos surgir as revisões sistemáticas de literatura, que como a nomenclatura esclarece, confere sistematicidade ao tradicional instrumento metodológico. Para Proença Júnior; Silva (2016), torna-se de natureza sistemática ao adotar critérios para a eliminação de vícios na consulta e uso de fontes não enviesadas. Tratando assim de uma "contribuição autoral que apresenta como resultado o estado da

literatura identificando as referências pertinentes a uma pesquisa". (PROENÇA JÚNIOR; SILVA, 2016, p. 234). Dentre as diversas qualidades da revisão sistemática de literatura, sua principal característica foca na obtenção de estudos prévios sobre determinada problemática, se propondo assim a uma meta-análise, promovendo estudos a partir de análises recuperadas. Essa exigência impactou negativamente seu uso enquanto metodologia para a proposta aqui pretendida devido a incipiente consolidação do campo das *digital humanities*, razão pela qual esta pesquisa não se baseará numa revisão sistemática. Todavia as demais qualidades de uma revisão sistemática, sobretudo àquela que tange clarificação dos critérios para sua reedição e política de minoração de vieses, servirão de parâmetros para a revisão de literatura promovida por esta pesquisa. Dessa forma o mapeamento da literatura se norteará por essa metodologia, apesar de não propô-la para este caso.

De toda sorte a metodologia de revisão que esta pesquisa pretende aplicar, mais se assemelha ao método das revisões narrativas, tidas como modelo tradicional por Segura-Muñoz; Takayanagui; Santos et al (2013) que traçam, a despeito das análises sobre revisão de literatura, um marco fundamental entre este e um modelo mais aprimorado como aquele classificado como sistemático. A distinção básica entre os dois e que nesta pesquisa nos fundamenta, se trata do caráter metodológico aprimorado de forma a ser passível de reprodução e sua objetividade em focar numa hipótese peculiar da pesquisa. Assim, baseamos nossa escolha pelo modelo tradicional/narrativo, porém criterioso com vistas a possibilitar sua reprodução.

A preocupação com o levantamento da literatura configura-se de vital importância devido à natureza fragmentada da produção da área. Como esse levantamento compôs os dados para o estudo bibliométrico, exigiu-se estratégias criteriosas de busca para obtenção dos estudos. De certo, como se trata de um campo com consolidação ainda incipiente, julgamos a contribuição bibliométrica factível para tecer uma imagem latente do corpo produtor de conhecimento nesta área tida como fragmentada em diversos outros campos do conhecimento. Nos fornecendo subsídios para, assim, possibilitar alguma identificação de fronteiras. De toda forma, a nosso ver, essas fronteiras puderam ser clarificadas pelas estratégias praticadas no âmbito da bibliometria.

A bibliometria, segundo Araújo (2006), constitui-se como técnica quantitativa e estatística para medição de índices de produção e disseminação do conhecimento científico. Surge de uma "necessidade de estudo e avaliação das atividades de produção e comunicação científica" (ARAÚJO, 2006, p.12). Para o autor sua atribuição serve à

ciência como o recenseamento serve à demografia, na qual o levantamento dos dados permite, senão uma imagem fidedigna do todo, uma possibilidade de identificar nuances que permitam conhecer e investigar características daquele conjunto. Para ele o que inicialmente voltava-se para uma medida de livros, atualmente agregou suportes outros como artigos e documentos, para voltar-se a uma utilização desses métodos quantitativos na busca por uma avaliação objetiva da produção científica. Ponto que configura-se como central dos estudos bibliométricos.

Apesar das três leis clássicas que fundamentam a bibliometria, outras teorias bibliométricas servem também de contribuição para aplicação desta técnica e sua consequente fundamentação empírica ao objeto desta pesquisa. A teoria epidêmica de Goffman; Newill (1964) que segundo Araújo (2006) explica a propagação de ideias dentro de uma determinada comunidade como um “fenômeno similar à transmissão de doenças infecciosas” (ARAÚJO, 2006, p.17). Para o autor essa teoria inspira a área mais importante da bibliometria que é a análise de citações.

Citação é um conjunto de uma ou mais referências bibliográficas que, incluídas em uma publicação, evidenciam elos entre indivíduos, instituições e áreas de pesquisa, visto que mostram o relacionamento de uma publicação com outra. A análise de citação pode ser definida como a parte da bibliometria que investiga as relações entre os documentos citantes e os documentos citados considerados como unidades de análise, no todo ou em suas diversas partes: autor, título, origem geográfica, ano e idioma de publicação, etc. (ARAÚJO, 2006, p. 18).

De certo temos a possibilidade de, ao buscar essas relações, mapear e identificar os atores, seus lugares de fala e a que assuntos se dedicam no universo analisado. Como pontua GRÁCIO (2016, p. 83) "a citação é tomada como indicador objetivo da comunicação científica, que evidencia as relações entre documentos e seus autores, tanto citante-citado, como citante-citante e citado-citado na visão do citante". Permitindo assim a percepção de um esboço que nos permitiu compreender um conhecimento novo sobre a análise proposta por esta pesquisa. Somamos a isso a constatação de que as citações se configuram um excepcional vetor de identificação de um campo já que

contribuem para o desenvolvimento da ciência, proveem o necessário reconhecimento de um cientista por seus colegas, estabelecem os direitos de propriedade da contribuição científica de um autor, constituem importantes fontes de informação, ajudam a julgar os hábitos de uso da informação e mostram a literatura que é indispensável para o trabalho dos cientistas. (Foresti, 1989, p.3 *apud* ARAÚJO, 2006, p.18).

Dessa forma exploramos essas possibilidades oferecidas pela análise de citações para identificar tais qualidades na literatura das humanidades digitais, sobretudo àquelas



que podiam permitir um despontar temático proveniente dessa produção. De toda maneira esperamos, a partir de então, possuir subsídios para promover um juízo de valor que tangencie o escopo da área objeto de estudo. Assim encontramos endosso no entendimento de GRÁCIO (2016, p. 84) da literatura científica como "a expressão mais objetiva da representação de um domínio científico, como resultado da atividade social de pesquisa que, neste contexto, enquanto estudo dessa literatura, se gabarita numa forma de análise do domínio".

Não obstante, os estudos bibliométricos se enriquecem de reflexões que se amparam em leis de origens lógicas e matemáticas, quais podemos citar, por exemplo, as leis de Zipf, de Bradford e de Lotka como as mais expoentes. A Lei de Lotka que segundo Araújo (2006) inspirou a lei dos quadrados inversos, na qual uma ampla proporção da literatura é produzida por um número pequeno de autores e que um grande número de pequenos autores, produz volume equiparável ao dos grandes produtores, se aplica às *digital humanities* com a peculiaridade de que em cada área do conhecimento esse ranqueamento pode ocorrer. Além de Lotka, a lei de Bradford, que recai sua análise nos periódicos científicos no que se convencionou chamar de lei de dispersão. Nela é possível julgar quais periódicos possuem uma relação próxima ou estreita com um determinado tema, no qual três zonas ilustram um pequeno número de periódicos altamente produtivos, um segundo número maior menos produtivo e um último grupo de periódicos muito maior, porém, bem menos produtivo.

Por fim a Lei de Zipf que, por teorizar a recorrência de palavras num determinado conteúdo, pode sustentar a análise das palavras-chave enquanto indicativo de um despontamento temático-conceitual. Dessa forma nos entusiasmos alcançar um mapa temático proveniente da análise bibliométrica de palavras-chave, tendo nessa lei uma contribuição para um embasamento teórico, pois descreve a relação entre palavras num determinado texto. Assim, segundo Araújo (2006), essa lei constata que ao listar as palavras que ocorrem num *corpus* textual em ordem de frequência, a posição de uma palavra na lista multiplicada por sua frequência é igual a uma constante. Essa formulação gabarita a análise do montante de palavras-chave a serviço do mapa de termos oriundos dessa literatura, tidos, portanto, como representação taxonômica latente do campo.

Todo o esforço de ordem bibliométrica, acaba por, conseqüentemente, imputar uma análise de domínio. A análise de domínio passou a ser um instrumento formal da Ciência da Informação por, grosso modo, captar subsídios para mensuração de um campo através de sinais oriundos de mapas bibliométricos. Para Hjørland (2002) os estudos

bibliométricos têm papel expoente nos estudos dessa natureza, pois é possível traçar onze abordagens responsáveis por cunhar os elementos balizadores mais legitimados da análise de domínio. Segundo ele essas orientações cumprem a função de nortear o trabalho de descrição de um domínio, passando sempre pela análise da literatura, tendo, assim, nos estudos bibliométricos, uma dessas onze abordagens. A importância da abordagem bibliométrica se dá por poder ser usada como ferramenta e método para análise de domínio de diversas formas, além de baseando-se nas análises de citação, popularizar a construção de mapas bibliométricos ou para visualização de áreas científicas. Assim a bibliometria é uma abordagem potente porque revela diversos detalhes e conexões reais entre indivíduos e documentos. Essas ligações “representam o reconhecimento do autor pela dependência entre artigos, pesquisadores, campos, abordagens e regiões geográficas”. (HJØRLAND, 2002, p. 432-433, tradução nossa)

Para Tennis (2012) a garantia literária do começo do século XX pode ser considerada uma vertente desse interesse, o que a nosso ver, motivou a formalização da área de estudo pela Biblioteconomia e Ciência da Informação. De acordo com o autor existem dois tipos de análise dessa natureza no qual seu tipo mais básico de análise de domínio parte da criação de mapas bibliométricos. Para ele trata-se de uma análise de domínio descritiva que evidencia a constituição de um domínio regida pelo interesse dos pesquisadores. Tendo, assim, “como heurística predominante para seu desenvolvimento a criação de mapas bibliométricos”. (TENNIS, 2012, p. 6). Dessa forma, podemos inferir que ao criar esforços para um estudo bibliométrico das *digital humanities* estamos também, latentemente, promovendo uma análise de domínio de ordem descritiva das humanidades digitais sob perspectiva da CI.

Para o Tennis (2012) esse processo descritivo passa por, necessariamente, três imperativos conceituais que viabilizam uma análise de domínio a saber, um; porque seguem métodos das Ciências Sociais, dois; por se embasarem na observação dos atores sociais, e três; por focarem num coletivo alheio às posturas subjetivas. Ainda segundo o autor, “a proposta de uma análise de domínio centra-se, fundamentalmente, não na definição de um domínio, mas em sua descrição” (TENNIS, 2012, p. 10). Dessa forma, para o autor, os debates sobre análise de domínio seguem refletindo sobre se a descrição se limitaria a uma forma bibliométrica ou se realmente já se concretiza numa ação ontologicamente descritiva, muito embora ambas de mesmo propósito. De toda maneira afirma que “não se está construindo um sistema de informação, mas simplesmente se descrevendo um domínio com vistas a pesquisas básicas” (TENNIS, 2012, p. 10).

Nomear um domínio é rotulá-lo, portanto ao utilizarmos o rótulo Religião, por exemplo, é importante detalharmos a sua extensão e as suas exclusões visando a comunicar o valor da atividade, ou seja, o valor da análise de domínio. O que se está considerando e o que não se está considerando na análise de Religião? Há que se considerar três coisas: nome, extensão e exclusões, as quais são consideradas áreas de modulação e este é o primeiro eixo a ser considerado ao especificar o escopo e alcance de um domínio quando se compromete com a análise de domínio (TENNIS, 2012, p. 8).

Como o que se propôs nesta pesquisa foi uma análise descritiva inicial e diagnóstica, devido às diversas características incipientes do campo, importa direcionar esta análise no sentido de limitar-se à identificação dos três elementos apontados pelo autor e representados pela definição do domínio: seu escopo, alcance e sua finalidade. Identificação esta proveniente de estudos de natureza bibliométrica, compreendendo, inclusive, um desvelar do despontamento temático fruto das palavras-chave da literatura.

## 1.4 Objetivo

Recuperar, mapear e analisar a literatura indexada por humidades digitais nas bases de dados generalistas, se utilizando de uma abordagem bibliométrica para mapeamento da área.

### 1.4.1 Objetivos Específicos

- ➔ Recuperar a produção acadêmica indexada pelo termo *digital humanities* e seus equivalentes através de uma revisão de literatura;
- ➔ Realizar uma análise das métricas da bibliometria baseado na literatura classificada como humanidades digitais;
- ➔ Construir e analisar grafos das relações bibliométricas e temáticas;
- ➔ Ensaiar uma análise de domínio do campo das humanidades digitais;

## 1.5 Metodologia

A viabilização da pesquisa perpassou, obrigatoriamente, o levantamento da literatura correspondente às humanidades digitais. Para implementar essa proposta,

pretendeu-se utilizar uma de revisão de literatura, inspirada nas qualidades de rigor metodológico aplicáveis à uma modalidade sistemática, para mapear a produção acadêmica da área de interesse como ponto de partida para o desenvolvimento da pesquisa. Dessa forma o objetivo da revisão foi, num primeiro momento, recuperar itens que abordassem as *digital humanities* nas bases de dados consolidadas como grandes repositórios globais de produção acadêmica. Dessa forma elegemos as bases Scopus e Web Of Science como os repositórios que atendem essa exigência, estando gabaritadas, segundo nossa proposta, por possuírem foco multidisciplinar e de amplitude ímpar. Essa preocupação se valida pela intenção de minorar algum tipo de viés implícito em bases dedicadas a alguma área predominante. A pesquisa se negou a utilizar bases peculiares, pois considera que a recorrência de periódicos próprios das áreas nessas bases deve interferir na análise, razão pela qual não foram consideradas como medida de minoração de vieses.

Com as bases devidamente eleitas, 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 ‘*Article title, abstract and keywords*’ 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

(“*Digital 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.

Os resultados recuperados nessa revisão foram extraídos nos formatos com potenciais possibilidades de análises por programas de computador, vedando o estudo

bibliométrico de análise manual, deliberando aos *softwares* informáticos a função de realizar tais análises métricas. A utilização desse ferramental enquanto metodologia de trabalho serviu à corroboração da própria praxiologia das HDs, uma vez que uma ação que outrora estaria delegada à uma análise morosa, de pouca abrangência e suscetivelmente enviesada, agora ganha em escala e apelo visual. Habilitando-se, assim, na revelação de relações e construção de conhecimentos outros que não possíveis de forma manual.

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. Nos 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.

Nesse sentido o VOSviewer é um programa de computador voltado a criação de mapas baseados em rede de dados para visualização e exploração desses dados. Pode ser usado principalmente para

construção de redes de dados de publicações científicas, periódicos acadêmicos, pesquisadores, instituições de pesquisa, países, palavras-chave ou termos. Os itens podem ser conectados nessas redes por coautoria, coocorrência, citação, acoplamento bibliográfico ou cocitação. Essas redes podem ser construídas a partir de dados extraídos diretamente de bases como Web Of Science, Scopus, PubMed, e outras com formatos compatíveis. (VAN ECK; WALTMAN, 2017, p. 3, tradução nossa)

O VOSviewer foi desenvolvido visando atender as demandas atuais de visualização de dados de origem bibliográfica, possibilitando uma reformulação nos instrumentos predecessores desta finalidade e proporcionando uma consonância deste campo de estudo com os padrões atuais de visualização de redes de informação. Dessa forma nos foi garantido pelo programa identificar inúmeras métricas e suas relações entre os links, sendo possível identificar visualmente a relação existente de, por exemplo, coautoria de dois autores ou de países, entre diversas outras. Assim, a proposta foi utilizar os recursos do VOSviewer para processar métricas bibliométricas analisáveis pelo *software* como, por exemplo:

Tabela 1: Sumário de possibilidades de análise para recursos bibliográficos

<b>Métrica</b>	<b>Análise</b>
Coautoria	Autores Organizações Países
Coocorrência	Palavras-chave
Citação	Referências Periódicos Autores Organizações Países
Acoplamento bibliográfico	Referências Periódicos Autores Organizações Países
Cocitação	Referências citadas Periódicos citados Autores citados

(Fonte: adaptado de Manual VOSviewer 1.6.6)

Através do VOSviewer foi possível analisar as métricas elencadas na tabela acima. Para cada uma dessas métricas e suas aplicações o programa permitiu a possibilidade de se obter o número de documentos identificados, sua quantidade de citações e um valor atribuído como 'peso'. Excetuando-se as análises de cocitação e coocorrência, esses elementos podem ser identificáveis pelo VOSviewer para as métricas de coautoria, citação e acoplamento bibliográfico. O programa atribui uma fórmula matemática que, relacionando quantidade de documentos frente seu número de citações, recebe uma escala de grandeza que os elenca em nível de relevância. A proposta da pesquisa foi apresentar essa ordem de grandeza baseada em relevância juntamente com os grafos de relacionamento que viabilizam a identificação visual das relações encontradas, além de interpretar essas disposições.

Segundo Grácio (2016) as métricas oriundas de estudos dessa natureza são altamente relevantes para seu contexto, como por exemplo, aquelas concernentes ao acoplamento bibliográfico que referem-se à relação entre dois artigos considerando o número de referências em comum citadas pelos dois artigos. Paralelamente a análise de cocitação mede a relação entre dois artigos com base no número de publicações em que estes são citados concomitantemente. Para VANZ; STUMPF (2010, p. 69) “a análise de coautoria é considerada uma das formas de se medir a colaboração científica e pode se referir a pesquisadores, instituições e países, estes dois últimos através da vinculação institucional dos autores”. Além dessas, não desprestigiamos a análise de citação, que visa identificar referências, autores, periódicos ou até organizações e países que, ao alcançarem uma patamar de citações considerável, passam a dispor de legitimidade acadêmica, além de influenciar a própria produção científica. Sendo assim possível, através dela, identificar qual o periódico ou autor mais citado numa determinada temática. Por fim temos a coocorrência de palavras que pretendeu-se aplicar às palavras-chave atribuídas aos artigos recuperados e também àquelas mineradas dos resumos.

## 2 HUMANIDADES DIGITAIS COMO EXPOENTE DE UM NOVO CONCEITO

Ao longo do tempo a relação do homem com o meio social e o natural experimentou mudanças contínuas e sem precedentes. Essas transformações impuseram uma nova forma de relacionar-se com a natureza e os atores sociais. A chegada da roda, da imprensa, do livro, da pólvora, do capital, do computador e tantos outros recursos emblemáticos da história, sinalizaram a descoberta potencializadora de uma nova configuração a ser experimentada pela sociedade ao longo do tempo e espaço.

No bojo dessas experimentações, pavimentou-se o caminho para conceitos que hoje balizam a sociedade contemporânea e permeiam os estudos sobre esse processo. Sendo esse movimento contemplado por diversos ângulos econômico, cultural, político e social fizeram despontar abordagens dedicadas a analisar tais fenômenos. De certo este movimento foi o responsável por construir temas maciçamente difundidos e conhecidos atualmente como, por exemplo, Globalização, Sociedade em Rede, Sociedade da Informação, Revolução Tecnológica, Era da Informação, entre tantos outros.

Do ponto de vista sócio-cultural contemporâneo uma das maiores reflexões desse processo se reverberou em torno do conceito de Sociedade em Rede, cunhado por Castells (2000). Para ele o conceito se configura, em linhas gerais, no paradigma pelo qual as relações sócio-econômico-culturais é alterado pela massificação das tecnologias da informação. Nesse novo modelo, que sintetiza bem o patamar atual da humanidade, essas relações passam a ser tidas e sentidas em interação, impactadas, basicamente, pelas tecnologias da informação e comunicação em escala global, influenciadas ainda pelo poderio econômico que engendrou-se num cenário pós-revolução industrial. Essa abordagem materializa uma percepção generalizada de um processo contemporâneo no qual a característica da atual revolução tecnológica não é a "centralidade de conhecimentos e informação, mas a aplicação desses conhecimentos e dispositivos de processamento da informação em um ciclo de realimentação cumulativo entre as inovações e seu uso" (CASTELLS, 2000, p. 69).

Paralelo a essa abordagem temos a ideia de sociedade da informação, que 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-lo. 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. Das formas de compra e leitura de um livro à consulta ao catálogo da biblioteca, a tecnologia pautou 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.

A CI apesar de dedicada enquanto campo de estudo a todo processo do trato informacional, acabou por atribuir espaço privilegiado para as reflexões em torno da massificação do tecnológico na área da informação. Algumas razões podem ser apontadas como balizadores históricos para essa especialização como a popularização dos sistemas automatizados de recuperação de informação e a virtualização dos serviços de bibliotecas e arquivos, e principalmente, pelo advento da informação em meio digital.

De certo, podemos cogitar que o digital vem promovendo uma revolução tal qual a industrial, por impactar a produção, o armazenamento, a recuperação, o acesso e a divulgação da informação e do conhecimento. Sua maleabilidade possibilita novos usos e apropriações elevando o patamar da relação homem *versus* informação. Todo esse caráter inédito até então, vem gerando reflexões e movimentos em torno do que se categoriza como revolução digital que tem se consolidado, especialmente no mundo corporativo, sob a alcunha de "Transformação Digital". Para Ross; Sennyey (2008) a revolução digital reformou todas as etapas de circulação da informação ao tornar o formato predominante, alterou a velocidade de criação da informação, da entrega à divulgação, além das necessidades e expectativas dos consumidores de informação.

---

<sup>3</sup> A autora destaca esse fenômeno enquanto contexto baseado numa observação do século XXI caracterizado por um desenvolvimento sem limites da ciência e da tecnologia, provocando novas necessidades, novas atitudes. Configura-se uma sociedade que busca o conhecimento e novos modelos que possibilitem interpretar e compreender o mundo: mentes que se antecipam, pessoas autônomas no pensar, sentir e fazer, com a capacidade de julgar situações, baseando-se em ideias e em fatos. (VARELA, 2005, p. 2)

Como esse não é um fenômeno restrito ao campo informacional, o grau de atividade desse processo pode ser constatado pela alteração que promoveu no universo corporativo, impondo a questão tecnológica, baseada no digital, no alvo das grandes corporações. Se no contexto informacional esse movimento estaria ligado à ideia superficial de "digitalização das coleções, da criação ao arquivamento como, provavelmente, a mais reconhecida manifestação onipresente da transformação digital" (DAIGLE, 2012, p. 251), no caso do mundo corporativo ela estaria ligada diretamente à questão da produtividade.

A abrangência desse movimento se fez experimentar não apenas no ambiente corporativo e informacional, mas também, não surpreendentemente, por todas as áreas do conhecimento. A produtividade, experienciada e almejada pelo mundo corporativo, se assentou sobre o universo acadêmico e seduziu as áreas a uma massificação da experimentação tecnológica amparada no digital em suas práticas tradicionais. Foi neste contexto que acadêmicos de diversas vertentes foram se dando noção de que a possibilidade de realizar tarefas, antes exclusivamente manuais, ganhavam praticidade fruto da disseminação dos recursos computacionais como preponderantes desse movimento global.

É a partir desse novo fenômeno, cuja a experimentação generalizada da automatização de fazeres no âmbito da pesquisa técnico-científica, baseando-se exclusivamente no suporte digital, vem sinalizar um novo prenúncio de nosso tempo. Essa sinalização prenuncia uma consolidação, nos poucos anos passados recentes, de um novo espectro de desenvolvimento prático-teórico como prelúdio do movimento que a partir do ano 2010 passou, formalmente, a ser nominado como humanidades digitais.

Assim a proposta deste capítulo é pensar e discutir as humanidades digitais como produto e sintoma das transformações tecnológicas cuja sociedade vem vislumbrando desde tempos mais remotos. Além disso, principalmente, pareando essas reflexões com aquilo que se identifica com interdisciplinaridade -característica esta que nos motivou a rastrear alguma relação de proximidade que nos corrobore a ideia de que, além da mera relação interdisciplinar, revelasse, também, uma justaposição.

## **2.1 Prelúdio à conceituação das humanidades digitais**

Considerar conceituar o que hoje se nomina como *Digital Humanities* implica, necessariamente, na observação de um movimento de ordem tecnológica nas atividades de pesquisa dos mais diversos campos das Ciências Humanas. Todavia esse é um movimento que se inicia bem antes do *status quo* das humanidades digitais (HDs), quando este ainda era tido sob a "ideia de 'bens comuns' de natureza metodológica para as técnicas computacionais compartilhadas entre as disciplinas das humanidades e ciências sociais estreitamente relacionadas" (MCCARTY, 2003, p. 1, tradução nossa). Até então, a penetração de um estatuto das tecnologias da informação e comunicação, davam a partida para um processo de consolidação do uso de recursos dessa natureza, basicamente pela possibilidade de bases de dados voltadas à recuperação da informação.

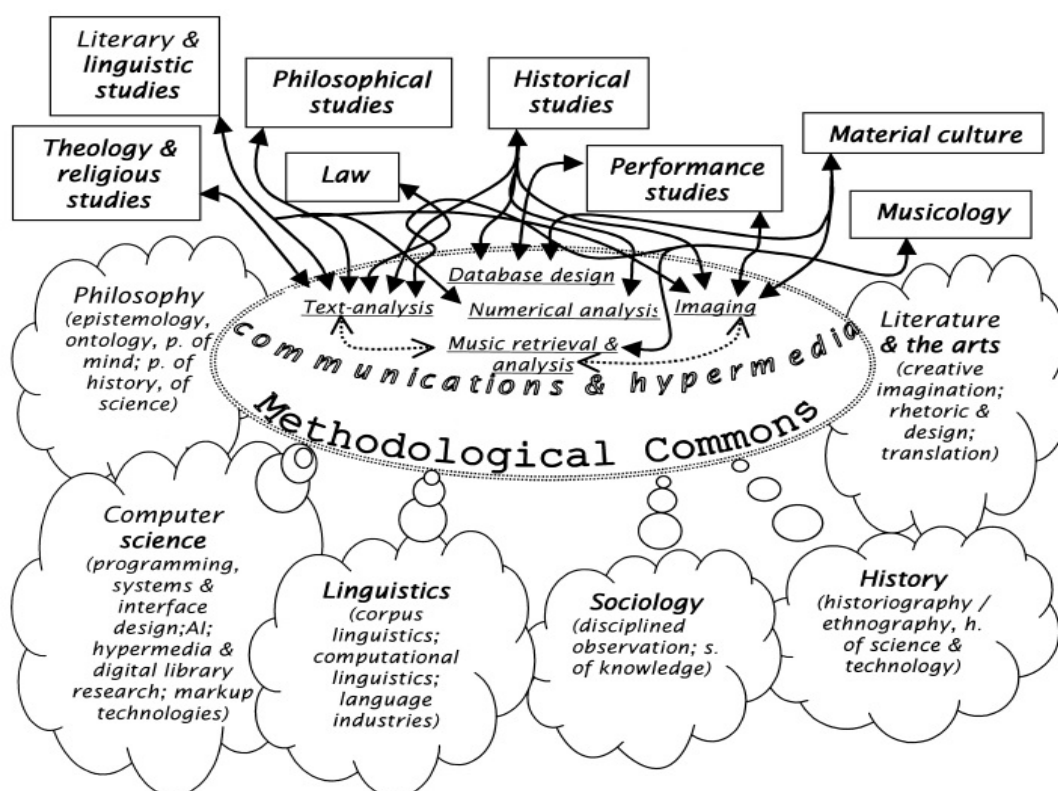


Figura 1: Esboço do mapa conceitual da Humanities Computing proposto por McCarty em 2003. (Fonte: MCCARTY, W. "Humanities Computing." In: Encyclopedia of Library and Information Science. New York: Marcel Dekker. 2003, p. 1225. Disponível em: <<https://www.mccarty.org.uk/essays/McCarty,%20Humanities%20computing.pdf>>. Acesso em: 04 de abril de 2019).

Para McCarty (2003) esse processo criou um intercâmbio de "bens comuns" de natureza metodológica que aproximou campos de pesquisa em torno da descoberta de novos métodos informáticos, motivando uma configuração de conexão entre campos

disciplinares (fig.1). Para ele essas modalidades de trocas mútuas configura a ideia de *Humanities Computing*

A atuação que observa essa troca de "importação/exportação" é a *Humanities Computing* em seu duplo papel como serviço colegiado para as disciplinas e como instrumento de pesquisa direcionada para investigar suas metodologias em evolução, a elaboração de novos métodos computacionais, abordagens, estudo de seus efeitos e provocação de suas implicações. (MCCARTY, 2003, p.1, tradução nossa).

Por essa razão alguns autores veem nas *Humanities Computing* o embrião das humanidades digitais, ou até mesmo como conceito precursor, por imprimir nesta abordagem as ideias fundamentais das HDs, como o robusto intercâmbio entre campos disciplinares e por aplicar essa observação ao fenômeno do advento tecnológico no desenvolvimento metodológico dessas áreas.

Todavia, ao nosso ver, esse conceito não necessariamente pode ser tido como sinônimo -ao nosso ver raso- das HDs como entendem alguns autores, no qual consideram que "as *digital humanities* são formalmente conhecidas como *Humanities Computing*" (TANG; CHENG; CHEN, 2017, p. 1). Para Svensson (2009), por exemplo, as reflexões em torno das quais se deram a *Humanities Computing* podem ser consideradas fundamentais para as HDs, mas num sentido transicional. Para ele, existem boas razões para pontuar o debate da *Humanities Computing* nas HDs por conceder uma aparente transição discursiva para humanidades digitais. Assim,

qualquer tentativa de mapeamento de um campo emergente pressupõe uma discussão de território e ambições disciplinares, e a *Humanities Computing* fornece um ponto de partida particularmente bom, já que é relativamente estabelecido e bem definido. [...] muitas das questões, considerações e parâmetros relevantes para a *Humanities Computing* também são relevantes para as Humanidades Digitais em geral. (SVENSSON, 2009, p.2, tradução nossa).

Dessa forma é possível já em McCarty (2003), mesmo muito antes de se cunhar o conceito formal de HDs, identificar parâmetros e considerações relevantes como aponta Svensson (2009) para o que, a partir de 2010, cunhou-se como *digital humanities*. McCarty (2003) já apontava o impacto da *hypermedia*, das bibliotecas digitais, do hipertexto, do algoritmo, das iniciativas de codificação de texto e etc nas práticas de ensino. Apesar disso, o autor reduz a importância do fenômeno nesse contexto pedagógico por julgar que ele se limitaria, através basicamente do *software*, a uma mudança de paradigma apenas para essas atribuições na qual o papel "ensino baseado em recursos", significaria "convergência de pedagogia e pesquisa no antigo modelo da

biblioteca, onde recursos singulares e relativamente imutáveis são separados de seus múltiplos usos altamente variáveis” (MCCARTY, 2003, p. 1228, tradução nossa). Assim, esse parece ser o ponto demarcado de distinção das abordagens no âmbito da *Humanities Computing (HC)* ante as *digital humanities*.

Com a exacerbação do cenário, pelas HC definido e determinado, a entrada em cena de um novo agente desse processo colocou as HC e as HDs numa relação justaposta como afirma Svensson (2009). Para o autor importa observar se a transição discursiva de um conceito para o outro é principalmente uma questão de "reembalagem", ou se o novo rótulo "também indica um escopo expandido, um novo foco ou uma relação diferente com o trabalho de computação nas humanidades tradicionais" (SVENSSON, 2009, p.7, tradução nossa).

Para Svensson (2009) há uma distinção inexorável entre os dois conceitos todavia sua relação de preeminência conceitual histórica levou os estudiosos associarem a *Humanities Computing* como passado e indistinta consolidação das humanidades digitais. Como tentativa de clarificar essa relação e suas respectivas implicações, o autor aponta quatro questões sobre fronteiras disciplinares e cultura epistêmica das *Humanities Computing*.

O primeiro ponto a ser levantado sinaliza como a abordagem é mantida em nível estritamente instrumental no tocante à tecnologia aplicada às humanidades; ou seja o foco da computação nas humanidades tradicionais não está inovando novas ferramentas, mas sim usando e desenvolvendo as existentes. Em segundo lugar há uma evidência que as abordagens em HC não prezam pelas relações interdisciplinares, apontando para uma percepção coletiva de um fenômeno a ser lido em particular, ou como explica o autor com uma orientação instrumental na qual não há motivo para questionar os "bens comuns metodológicos" como um foco interdisciplinar valioso e um sentimento colaborativo produtivo (SVENSSON, 2009, p. 9). Assim, segundo ele, esse enfoque afeta a forma como a HC opera e se relaciona com outras disciplinas sendo a aplicação mais séria do vínculo predominantemente metodológico que pode não integrar muitas das questões específicas que estão no cerne dessas disciplinas.

Um terceiro ponto que marca essa distinção está no acentuado foco textual das HC, transparecendo uma relegação à segundo plano dos variados suportes tratados no âmbito das *digital humanities*. Assim, dada a história e preocupações primárias do campo da HC, bem como a orientação textual de muitas das humanidades, "não é surpreendente seu predominante foco textual" (SVENSSON, 2009, p.10, tradução nossa) a despeito do

que parece supor um rechaço aos diversos outros instrumentos de métodos e análises atualmente abarcados nas *digital humanities*.

A quarta e última deliberação de distinção entre as *Humanities Computing* e as humanidades digitais se apresenta relacionado aos suportes no qual Svensson (2009), identificando em McCarty (2005) que o ponto de distinção mais direta está relacionada ao escopo de suportes. De certo, segundo o autor, distingue-se entre quatro tipos de dados de abordagem metodológica comum da HC: "texto, imagem, número e som, sendo característica de um modelo no qual os materiais de origem e as abordagens das disciplinas são reduzidos desses quatro tipos de dados a uma finita configuração de manipulação dos mesmos" (McCarty, 2005 *apud* SVENSSON, 2009, p.10-11, tradução nossa).

Com bases nessas abordagens, podemos perceber que a figura do digital esteve sempre preterida, em segundo plano ou ainda sequer vislumbrada nas reflexões acerca das *Humanities Computing*. Esse é ao nosso ver a distinção mestra que viabiliza as abordagens "finitas" como aquelas vistas anteriormente no escopo da HC. Se devendo, dessa forma, ao suporte digital as inovações metodológicas mais atuais, sinalizando assim, as HC como desalento ao "infinito" escopo de atuação das *digital humanities* baseada no suporte digital. De certo,

[...]pode-se argumentar que a *Humanities Computing* está principalmente interessada em textos digitalizados (ou em alguns casos, sites históricos digitalizados, etc.) e não em material que é nativamente digital. O nativo digital incluiu jogos de computador, blogs, mundo virtual, redes sociais, coleções de e-mail, sites, imagens de vigilância, filmes e arte digital. A maioria desses "objetos" são estudados e analisados em diferentes tipos de novos meios de comunicação e configurações, e para mim, esta é uma interessante zona intermediária. A *Humanities Computing* estaria interessada em agregar mais com novos especialistas? Certamente há necessidade de ferramentas bem desenvolvidas para estudar a vida e a cultura on-line. (SVENSSON, 2009, p.11, tradução nossa).

Se podemos considerar que a *Humanities Computing* serviu de transição para o que hoje nominamos de humanidades digitais, correspondendo em grande parte às problemáticas das HDs, porém não abrangendo todo seu escopo, podemos afirmar que a entrada em cena dos objetos digitais, sobretudo aqueles nativamente digitais, foram responsáveis por alçar não apenas o debate, bem como todo o universo metodológico, ao patamar das *digital humanities*.

## 2.2 Objetos digitais e o não-finito: da maleabilidade à uma ontologia

O movimento de automatização de praticamente toda a interação humana teve na digitalização seu maior expoente, pois através desse processo de transformação pôde conferir diversas novas formas de transmissão, armazenamento e acesso aos registros informacionais. Essa engrenagem, que anteriormente tinha sob sua gerência itens físicos por meio de computador como os sistemas de recuperação de informações, hoje é capaz de produzir em ambiente totalmente não físico além de, também, igualmente gerenciá-lo. Assim, dessa forma, ao produto desse processo denominamos *born-digital*<sup>4</sup>.

Como '*born-digital*' entenda-se exclusivamente ao dado produzido sob a égide do digital e sem ela inexistente. Como define MUHANNA (2018) o *born-digital* significa, em alguns casos, “referir-se a registros ou arquivos que não têm contrapartida física: coisas que começaram a existir enquanto entidades digitais em vez de prepostos análogos convertidos em forma digital” (p. 2, tradução nossa). Pelo que vimos até agora, é nesta entidade que se encontra a base para constituição de uma nova área -das *digital humanities*- a partir do avanço das reflexões propostas pela *Humanities Computing*.

O *born-digital* sinaliza a instauração de um elemento capaz de reordenar toda a cadeia da lógica informacional, uma vez que sua contribuição "não é apenas a introdução de um novo habitat" para o trabalho nas humanidades, mas sim o "oferecimento de um conjunto de novos mecanismos para obter, visualizar e manipular esses dados". Ao fazê-lo, as tecnologias digitais "contribuem com a dimensionalidade, eficiência, acessibilidade e interatividade para modelos estruturais que foram, por milênios, utilizados." (MUHANNA, 2018, p.1, tradução nossa).

A categoria que representa esse conceito, cuja produção se dá por meio exclusivamente digital, marca uma nova etapa de um processo que, na verdade, se iniciou no movimento da digitalização em massa. Vários objetos cujo suporte físico sofreu um processo de migração para o formato digital, auxiliaram na construção do que hoje se constituiu num processo irreversível. Esse processo foi responsável por reafirmar a importância ou valor dos suportes originais, tidos como as fontes primárias. Dessa forma a característica que assola os *born-digital* é, em primeira medida, sua peculiaridade de ter

---

<sup>4</sup> O termo *Born-Digital* (em tradução literal 'Nascido Digital') em sua conceituação original em língua inglesa refere-se ao suporte produzido em meio digital e apenas através deste acessível, porém sua tradução para Português cunhou-se em torno dos indivíduos "Nascidos Digitais" que, em linhas gerais trata-se daquele que nasceu e cresceu com as tecnologias digitais presentes em sua vivência. Em inglês essa mesma definição se apresenta em torno do conceito de 'Digital Natives'. Já em Português o equivalente ao '*born digital*', enquanto sentido pretendido desta pesquisa, responde como 'Nato Digital', todavia esse é um conceito consolidado e restrito ao contexto da disciplina arquivística, tangendo os documentos da esfera pública produzidos digitalmente. Por essa razão elegemos nos referir neste texto ao termo original em língua inglesa '*Born Digital*'.

no impalpável a qualidade ímpar da fonte primária. Ou seja, o que no passado possibilitava que apenas a informação migrasse do suporte físico para o digital, atualmente não carece de tamanha exigência de validação, pois o digital já desfruta de credibilidade enquanto forma primária do registro.

Para Kirschenbaum (2013) esse processo obrigou uma reflexão em torno do produto da digitalização fazendo com que instituições manifestassem a importância do documento primário ante a "onda generalizante" de digitalização das coleções. Segundo ele "a fundamentalidade dos registros primários necessitou ser ressaltada, em 1995, em meio à primeira onda generalizada de digitalização, pela *Modern Language Association* que emitiu uma "declaração sobre o significado dos documentos primários" (KIRSCHENBAUM, 2013, p.1, tradução nossa). Nessa declaração a instituição afirmava a importância de manter livros e outros artefatos físicos, mesmo depois de terem sido microfilmados ou digitalizados para consumo geral.

Todavia, atualmente, a colocação impositiva fruto da massificação do que conceituamos por *born-digital* é o responsável por alterar o paradigma, posto que a fonte primária doravante se constituirá também num suporte imaterial. Assim, hoje, a presunção de um "'registro primário' não pode mais ser assumida como coincidente com a de um "objeto físico". Textos eletrônicos, arquivos, feeds e transmissões de todos os tipos também são, indiscutivelmente, registros primários". (KIRSCHENBAUM, 2013, p.1, tradução nossa).

De toda forma, vemos nesta imposição, o estabelecimento de um modelo que convergiu para a preponderância de um formato. Este modelo é abordado sob uma perspectiva filosófica por HUI (2012) que vê nesta instituição uma "'ontologia do digital' que consiste em dois conceitos principais: primeiro, que os bits são a representação atômica do estado da informação; e, segundo, que o estado temporal da evolução é um processo de informação digital" (p. 381, tradução nossa). Encontramos nessa abordagem a fundamentação para considerarmos a importância deste tema independentemente de sua característica *born-digital*. Todavia sua segunda categorização deste conceito, em que atribui ao estado temporal da evolução enquanto um processo de informação, identifica, ao nosso ver, a questão da informação gerada exclusivamente em meio digital.

À vista disso sua abordagem filosófica nos fornece a ideia de 'dado', embutida na de objetos digitais, que nos remete a uma percepção de "datatificação" dos objetos e a "objetificação" dos dados. Pois



[...]devemos reconhecer que, desde 1946, a palavra "dados" teve um significado adicional: "informação transmissível e armazenável por computador". Este segundo sentido de "dados" sugere uma reconsideração da filosofia dos objetos, já que o dado não pode mais ser tomado como dado sensível ou um modo de estar junto do homem e da natureza; Em vez disso, é preciso reconhecer sua transformação material. O significado da nova técnica de processamento de dados que chamamos de digital não é apenas aquilo que podemos processar em grandes quantidades de dados com auxílio de computadores, mas também o que o sistema pode estabelecer conexões e formar uma rede de dados que se estende de plataforma em plataforma, banco de dados para base de dados. (HUI, 2012, p.388, tradução nossa).

Essa questão conduz a análise para a convergência das atuações dos indivíduos íntimos às computações e aqueles demais quais presenciaram a migração de seus objetos de análise para o digital. Esse cruzamento, que tem em práticas de operação das informações digitais como XML, Web semântica, padrões de metadados etc, configura uma terceira fase. Para Hui (2002) a contraposição de objetos naturais e digitais, o papel do 'dado' e o estabelecimento de redes fruto do processo do digital pontuam seu debate filosófico em torno dos objetos digitais dentro de 3 fases: *objects, data, and networks*.



Figura 2: Fases na abordagem filosófica dos objetos digitais proposta por Hui (2002). Fonte: do autor.

Assim, apesar de se configuram fases distintas e a análise não se limitar a elas, objetos digitais, confrontados em suas propriedades técnicas e naturais, constituem-se dados e, estes, fontes para estabelecimentos de redes. Dessa forma, se a investigação de objetos naturais estiver relacionada com a "dialética de assunto e substância", e a "investigação de objetos técnicos está em causa com a relacionalidade entre o objeto e o meio", então a investigação de objetos digitais deve obter uma nova direção, potencializando estas duas investigações ainda mais (HUI, 2002, p. 390, tradução nossa).

Essas 3 fases, além de, ao nosso ver, credenciarem os objetos digitais a um patamar "não finito",<sup>5</sup> podem encontrar reverberação em Yamaoka; Gauthier (2013) por, segundos estes, terem os objetos digitais três classes baseadas em Thibodeau (2002). Sendo essas classes, para eles, objetos de ordem física, lógica e conceitual. Em sua ordem

<sup>5</sup> Cf. Svensson, 2009, p.10-11

física se aproxima da primeira fase de Hui (2002) que defende a resolução de objetos naturais ante os técnicos, cuja ideia predominante é sua qualidade física; em sua ordem lógica se aproxima de sua capacidade de reserva de 'dados', que conseqüentemente, se relaciona com a ordem conceitual. Em sua terceira classe, produz, assim, sentido a partir dos dados e possibilitando então o estabelecimento de redes, que se configura enquanto última fase proposta por Hui (2002).

Ainda para Yamaoka; Gauthier (2013) é possível propor, baseado em Kallinikos; Aaltonen; Marton (2010), uma teoria para os objetos digitais. Esta teoria estaria baseada em 4 atributos que lhes confeririam um perfil próprio: editabilidade, interatividade, reprogramabilidade e perenidade de fronteiras. Sua editabilidade garantiria sua maleabilidade, sendo passível de transformação ilimitada; sua interatividade garantiria total integração com o indivíduo em leitura, manuseio etc; sua reprogramabilidade nos fornece a ideia de interoperabilidade que capacita ao digital sua integração não apenas com o indivíduo humano, mas, principalmente, com outros elementos digitais. Por fim sua perenidade de fronteiras permite a alta e eficaz difusão de seus elementos.

Por conseguinte as características do suporte digital foram responsáveis por pavimentar o caminho a uma nova estrutura de trabalho no âmbito das ciências, reprogramando totalmente os métodos e instrumentos de pesquisa sobretudo nas disciplinas não íntimas, até então, ao fenômeno das TICs. Conseqüentemente esse processo, ao nosso ver, é parte indissociável ou mesmo condição *sine qua non* para o que esta pesquisa propôs-se investigar. Dessa forma subscrevemos o que Evens (2012) classifica enquanto uma ontologia do digital, que propõe uma reflexão a respeito do sentido abrangente deste atributo, como aquilo que torna possível suas múltiplas existências. Nos apropriando dessa concepção para pactuar através dela a transição para o que culminou na concepção de *digital humanities*.

O digital tem uma ontologia, uma forma de ser, produtos e processos gerados através de tecnologias digitais trazem vestígios desta forma de ser. Não obstante a aura tóxica do determinismo tecnológico em torno desta afirmação proponha que a arte digital e a mídia, os procedimentos digitais e até usuários de tecnologias digitais se comportem como essas tecnologias, para compartilhar sua ontologia. **A característica do digital é tornar a abstração materialmente operacional, trazer essa abstração para o concreto, sem que ele deixe de ser abstrato.** Principalmente, incorporando abstração, as pessoas, objetos e eventos de cultura digital em conexão e envolvimento mútuo com as tecnologias digitais no coração dessa cultura. (EVENS, 2012, p. 3, grifo nosso, tradução nossa).

De certo consideramos que essa ontologia, expoente da predominância do instituto do digital, foi o ponto de partida, o gene por natureza, e cerne por consequência, do que se desenvolve como *digital humanities*. Dessa forma responsabilizando-se por reunir em torno de uma incorporação abstração, pessoas, objetos, eventos e outros em conexão mútua, influenciando seu comportamento e refletindo essa lógica. Assim revela-se um comportamento coletivo ante ao determinismo tecnológico que impulsiona a agregação de métodos, indivíduos e reflexões em torno das humanidades digitais.

### 2.3 As humanidades digitais além de um esboço conceitual

Podemos apontar que a área das HDs se configuram em termos práticos ainda antes de sua rotulação como tal, sendo constituída a partir de questionamentos e reflexões de tradições arraigadas neste contexto específico. Mesmo antes de sua "fundação" formal enquanto campo, a área já era percebida e delineada por teóricos envolvidos no processo tecnológico que arrebatou as metodologias tradicionais para um diálogo compulsório com os computadores, *softwares* e indivíduos. Para Dalbello (2011) uma constatação fundamental foi perceber a tecnologia digital como ferramenta para "o programa de humanidades" como centro da consideração dessas áreas.

Seguiu-se daí o que a autora pontua como um gênero de discurso caracterizado por um repertório de temas que oferecem explicações regulatórias e reflete sobre os pressupostos subjacentes de seus objetivos da prática, configurando-se no "estudo da transformação da escrita, do pensamento e do ensinamento nos campos humanísticos" promovidos neste contexto (DALBELLO, 2011, p. 1-2, tradução nossa).

Pontua-se, segundo Dalbello (2011), que esse gênero de discurso surge ainda nas reflexões mais passadas no âmbito das *Humanities Computing*, que apesar de não configurarem o exato sinônimo da abordagem de um fenômeno que em grande escala se trata de um mesmo, mas que complementa-se ao primeiro a partir da entrada em cena do instituto do digital, compartilham um núcleo comum. Assim o que, emblematicamente, identificou-se na publicação do Index Thomisticus<sup>6</sup> como instauração das reflexões em

---

<sup>6</sup> O emprego inicial mais significativo de recursos computacionais nas humanidades foi o Índice Thomisticus, uma compilação das obras de São Tomás de Aquino iniciadas pelo Padre Roberto Busa. "O índice tornou-se a base para a publicação do banco de dados de versões dessa compilação e é considerado o primeiro projeto de texto eletrônico nas humanidades. O projeto começou como uma parceria governamental-acadêmico-industrial, lançando a primeira geração de cálculos digitais em larga escala da

torno da *Humanities Computing* enquanto constituição de bases de dados voltadas à aplicação de ferramentas informáticas em texto eletrônico, atualmente se desdobra na ideia de que

As humanidades digitais hoje abrangem uma gama de práticas e produtos acadêmicos, incluindo *corpus* lingüísticos, arquivos digitais interativos e projetos de edição. Alguns desses *corpus* de dados digitais tornaram-se recursos essenciais de suas disciplinas. Decorrente de esforços para construir coleções de materiais de fontes primárias gerenciando e organizando projetos de edição acadêmica em grande escala, eles também forçaram especialistas das humanidades a se tornarem inovadores tecnológicos. Também forneceram um foco para pensar possibilidades de texto eletrônico, já que as coleções emergentes foram consideradas criadores através das lentes da teoria da narrativa, da estética e da política das redes. (DALBELLO, 2011, p.3, tradução nossa).

Para GARDINE; MUSTO (2015) há uma mudança crucial na maneira pela qual os acessos digitais "preservam, agregam e desagregam, apresentam, privilegiam e refletem de volta a erudição que pode deixar categorias antigas para trás e mudar a maneira como até os humanistas pensam, pesquisam, escrevem, publicam e interagem com suas próprias comunidades" (p. 283, tradução nossa). Assim, o foco nessa mudança sugerida pelos autores, a um regressar à história, por isso identificamos em Dalbello (2011) a construção de uma narrativa, 'um gênero de discurso', um mesmo contexto a ser observado por pares que passam a se congregam em torno de um mesmo núcleo.

A construção de uma narrativa produzida "em torno da interseção entre as humanidades e o digital criou um ambiente no qual as humanidades se tornaram sujeitas a novas abordagens que levantaram questões sobre a natureza das humanidades". (GARDINE; MUSTO, 2015, p. 294-298). Sendo expoente dessa consideração que identificamos como elemento genético do campo e fundamental para interpretação de seu domínio está o manifesto de promulgação das *digital humanities*. Construída formalmente a partir de reflexões realizadas no âmbito THATCamp Paris em 2010, indivíduos oriundos das mais diversas matizes acadêmicas alcançaram o rótulo que significaria as reflexões em torno da conceitualização deste fenômeno.

Segundo o MANIFESTO (2011, não paginado) existem basicamente três pontuações fundamentais em sua definição, sendo esses acerca da "questão digital, da mobilização de instrumentos e perspectivas singulares do mundo digital restritos ao universo das disciplinas humanas e da transdisciplina" fruto das perspectivas heurísticas

---

IBM com máquinas para o trabalho de pesquisa de quase duas décadas antes da indústria de computação a começar expandir seu alcance em massa". (DALBELLO, 2011, p. 3, radução nossa).

advindas desse fenômeno. Dessa forma a "opção da sociedade pelo digital" seria a responsável pela reconfiguração do que identificamos como uma "alteração e questionamento às condições de produção e divulgação dos conhecimentos". A essa percepção identificamos enquanto uma ontologia que, a nosso ver, define a reconfiguração e potencialização da limitação do suporte tradicional a uma escala de incomensurabilidade antes não observada.

Em segundo lugar, no que tange a mobilização de instrumentos e perspectivas, defendemos que se corrobora a ideia de um gênero de discurso responsável pela criação de uma narrativa pseudo inovadora, que congrega em torno de si agentes, que cientes da compulsoriedade das técnicas, assumem os produtos desse fenômeno apenas enquanto aditivo em suas práticas metodológicas. Entendemos esse comportamento sob a forma passiva, uma vez que grande parte da área das humanidades não possuía em seu arcabouço nenhuma característica que as capacitariam a interagir com as novas tecnologias enquanto relação mútua antes da concepção da ideia de *digital humanities*. Essa conjectura nos serve para delimitar a existência de disciplinas que estabeleciam relação mutual anteriormente à configuração das HDs. Assim defendemos que, mesmo pertencente ao domínio das humanidades, áreas como a Ciência da Informação estabeleciam previamente à conjectura de *digital humanities* um vínculo recíproco já fortemente problematizado, servindo, portanto, de constatação à essa argumentação.

Em sua terceira pontuação, enquanto certificação de uma "transdisciplina" temos a constatação de um novo campo alheio às humanidades, porém integrado por essas disciplinas. Dessa forma significa dizer que, da mesma forma que existe uma relação mutual distinta no que tange a mobilização de instrumentos e perspectivas, também aqui constatamos que essa transdisciplina compreende a ideia de inauguração de um novo espaço reflexivo qualificado porém, também, compreendido por essa mesma lógica na qual existam relações mútuas. Relações essas que já denunciavam que determinadas disciplinas trouxessem em seu bojo uma espécie de justaposição que, configurada pela relação mutual, corroboraria a tese de distinção entre as disciplinas das humanidades face à constituição das *digital humanities*. Assim, dessa forma, advogamos a tese de que compreendido um núcleo comum configurado pela HDs, existem gradações peculiares dos efeitos que estas promoveram no interior das humanidades. Competindo, portanto, enquanto "fruto das perspectivas heurísticas advindas desse fenômeno" (MANIFESTO, 2011, não paginado) uma participação mais ou menos eloquente das disciplinas das humanidades. Por essa complexidade importa objetivar que as disciplinas humanas

participaram deste fenômeno desde sua concepção enquanto *humanities computing*, assumindo relevância distinta e viabilizando a incorporação de temáticas próprias das tecnologias<sup>7</sup> em suas abordagens em escalas consideráveis.

O Manifesto (2011) também indica que as "experimentações no domínio do digital em Ciências humanas e sociais nos últimos cinquenta anos" configuraram a emergência da área. Posto que, neste contexto, denuncia um despontamento da "emergência, de mais recentemente, centros das HDs, que são todos, atualmente, protótipos ou lugares de aplicação específica de uma abordagem das *digital humanities*" (MANIFESTO, 2011, não paginado). Dessa forma ao fazer menção aos novos espaços configurados vê-se um apelo praxiológico desgarrado das disciplinas em particular, sugerindo levar a este novo campo a experiência destas disciplinas a um núcleo coletivo. Mais uma vez, defendemos que este movimento não pode ser compreendido de forma uniformizada, haja vista que sustentamos que determinadas abordagens específicas dessas experimentações promoveram uma conjunção que, a despeito do movimento concernente das HDs, causou mudanças irreversíveis no seio de determinadas áreas. Assim, áreas outras experimentaram não apenas uma transformação predominantemente metodológica, mas também técnica, estrutural e organizacional. Por certo o que o documento indica como "aplicação específica de uma abordagem das *digital humanities*" (MANIFESTO, 2011, não paginado) defendemos enquanto propriedade de considerar o movimento das HDs de forma bastante peculiar face à Ciência da Informação, por ser esta um caso típico do que nos referimos como promotora de uma conjunção, por ter esta também sido impactada sob o ponto de vista técnico, estrutural e organizacional.

Uma outra indicação do (MANIFESTO, 2011, não paginado) dá conta de que o "digital induz uma presença mais forte dos aspectos técnicos e econômicos na pesquisa", pois naturalmente faz coincidir as questões de desigualdade uma vez que o fator econômico assume determinismo frente à necessidade de recursos para implementação

---

<sup>7</sup> Podemos tomar como exemplo a questão da migração das bibliotecas para suas versões virtuais e digitais. A área da informação teve a implementação dessas transformações o que compulsoriamente levou para a respectiva área problemas próprias da informática. Questões dessa ordem foram categorizadas sob novas percepções tal como o Serviço de Referência Digital que, segundo Alves; Vidotti (2006) procurou transpor para o ambiente digital, o Serviço de Referência e Informação realizado tradicionalmente em bibliotecas, com a ajuda de recursos apropriados, e utilizando-se dos benefícios das tecnologias de informação e comunicação (TIC's) para otimizar seus serviços e atender às necessidades informacionais dos usuários da unidade de informação destacando que as atividades desenvolvidas pelo Serviço de Referência foram sendo alteradas e incrementadas em virtude do aparecimento de novas tecnologias. Deste modo, como uma alternativa aos meios e necessidades atuais surge o SRID, realizado por bibliotecas digitais, disponibilizado via World Wide Web e que possibilita consulta a um profissional da informação com respostas mais elaboradas aos questionamentos do público potencial. (ALVES; VIDOTTI, 2006, p.2)

computacional. Por estar passível de reproduzir e exacerbar as desigualdades é urgente questionar a alocação de recursos de natureza pública, além de sensibilizar o fomento de natureza privada ao ensino e pesquisa. Mais uma vez se faz pesar as características técnica, estrutural e organizacional, pois sobre elas recaem ainda mais a relevância de obtenção de recursos econômicos. Desta forma a insuficiência de financiamento pode determinar "que esta obrigação é uma oportunidade para fazer evoluir o trabalho coletivo", fazendo do propósito coletivo um capacitante do compartilhamento de recursos técnicos como *softwares* livres, programas de computador em código aberto e ações na esfera dos dados abertos conectados.

O Manifesto também indica "que existem diversos métodos provados, conhecidos e compartilhados desigualmente" (MANIFESTO, 2011, não paginado) que, por conseguinte, nos obriga a refletir sobre como esses novos métodos receberão ressonância em seus públicos-alvo. Com a produção de conhecimento estando cada vez mais dependente de programas de computador capazes de realizar análises inconcebíveis a capacidade humana, deve-se pontuar a capacidade destes operarem os *hardwares* e *softwares* instrumentos desta nova metodologia. Desta maneira a capacidade individual deve ser colocada no centro da questão, pois significa dizer que o conhecimento não evoluirá à velocidade proposta pelas HDs se as pesquisas e os pesquisadores não usufruírem de competências apropriadas. Essa é uma problemática que já outrora encontra ressonância na Ciência da Informação sob esse viés, assim, como defendido, identificamos questões que já podem compreender uma antecipação por determinadas áreas como neste caso pela CI. Decerto, temos na questão das competências individuais, mais um indicativo da relação mútua que promoveu além de transformações, uma mutação que pode ser categorizada como uma antecipação à concepção de *digital humanities*.

Em uma última indicação, o Manifesto delinea aquela que deve corresponder, ao nosso ver, como fulcro daquilo que se discute, assim as "múltiplas comunidades específicas, oriundas de interesses por diversas práticas, instrumentos ou objetos transversais" (MANIFESTO, 2011, não paginado) exponenciam as *digital humanities* e legitimam sua existência. Independentemente das nuances que eventuais abordagens possam imprimir no ou sobre o campo, é neste atributo que congrega aquilo que de mais autêntico representa a área particularmente sinalizados por seus "instrumentos ou objetos transversais".

Dessa forma a observação desse fenômeno, seguida de sua consecutiva observação narrativa, nos sugere propor um diagnóstico desse cenário que constata-se como base, em linhas gerais, em três principais alicerces daquilo que pavimentou o percurso para as HDs. Portanto, prescrevemos enquanto condição *sine qua non* para a construção do campo das *digital humanities* o tripé das atividades de pesquisa, os objetos destas e as técnicas à estes aplicados. Dessa forma esquematizamos:

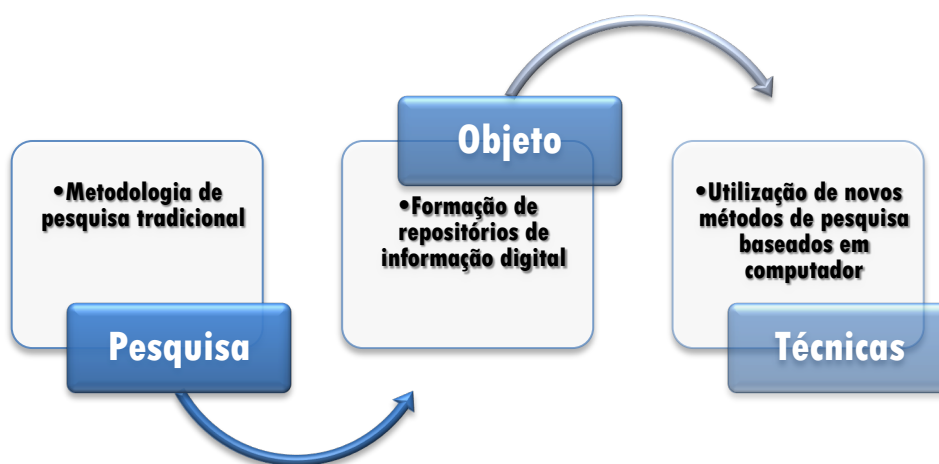


Figura 3: Esquema proposto à concepção das humanidades digitais (Fonte: elaborado pelo autor).

Isto posto identificamos que, por atividades de pesquisa, compreenda-se toda a ação brutalmente impactada pelo movimento computacional, partindo de seu modelo 'tradicional'. Essa atividade metodológica confronta-se, nesse primeiro momento, com o objeto de pesquisa que, agora sob uma massificação enquanto objetos digitais, acumulados em grandes repositórios de informação, potencializa o fenômeno do processamento computacional desses registros. O processo que conduziu ao encontro esses dois elementos -objeto e método- criou bases para uma possibilidade inovadora na qual a análise não mais se recairia sobre o suporte, mas sim sobre conteúdo, ou seja seu atributo informacional. Assim decreta-se a limitação do suporte e abre-se margem ao caráter heurístico das análises do suporte digital por novas ferramentas informáticas. A nosso ver esse é um momento prévio que baliza e sustenta a compreensão das humanidades digitais.

Assim vemos surgir não apenas o vislumbrar de novas experimentações metodológicas, mas, concomitantemente, reflexões sobre este fenômeno. Para SULA (2013, p.4, tradução nossa) as HDs enfocam tanto "a aplicação da tecnologia de computação às investigações humanistas quanto as reflexões humanistas sobre o



significado dessa tecnologia". Essa nova configuração confere às *Digital Humanities* o contorno de uma nova disciplina, pois conjuntamente com essa constação observa-se igualmente a visão tecnocêntrica dos personagens advindos dessa área correlata. Como observa GARDINE; MUSTO (2015, p.319, tradução nossa) “o forte contraste entre as duas abordagens demonstra a natureza contestada do termo, e talvez isso seja o resultado do fato de que enquanto humanistas e cientistas da computação estão em diálogo, cada um com sua própria perspectiva distinta, o elemento digital da definição está subjacente a ambos”. Para Russel (2011) esse é um processo que vem criando um corpo de indivíduos que se identifica como humanista digital, ou seja indivíduos que além de participarem ativamente das inovações do novo campo, também se dedicam a interpretá-lo. Assim, ao nosso ver, tais características convergem esses fenômenos no sentido de designar uma transdisciplina que congrega em torno de si métodos compartilhados, objetos comuns entre agentes distintos e que vem se identificando como fabricantes de um mesmo produto, pertencente a um escopo comum. Com base nisto, sugerimos esquematizar as HDs na seguinte estrutura:



Figura 4: Proposta lógico-conceitual para as humanidades digitais (Fonte: elaborado pelo autor).

Portanto advogamos que a HDs partem de pressupostos configurados enquanto uma ontologia do digital que confere ao suporte uma maleabilidade, responsável por possibilitar as quase infinitas novas experimentações. Com a potencialidade que esse recurso proporciona, identificamos uma congregação de agentes das mais diversas origens disciplinares para compor experiências provenientes das técnicas computacionais. Essa ação caracteriza propriamente as *digital humanities* e a abordagem

auto-reflexiva que fundam, cunham e fundamentam esse novo conceito. Na visão de Russel (2011) estamos diante de um termo que busca entender o impacto e o relacionamento das tecnologias de computação na tarefa de pesquisadores nas humanidades e que oferece novas e excitantes possibilidades para o desenvolvimento da pesquisa e do ensino em humanidades. A autora elenca que, de maneira geral, os objetivos das HDs são a "criação de bancos de dados com recursos digitais relevantes para as humanidades; o desenvolvimento de metodologias que permitam gerar novos elementos derivados desses dados e a geração de pesquisa e conhecimento para aumentar nossa compreensão nas humanidades" (RUSSEL, 2011, p. 3, tradução nossa)

Todas essas constatações entorno do que vem a considerar-se como *digital humanities* impõem uma configuração ainda mais complexa, que ganha forma acerca da reflexão sobre o próprio campo e seu escopo. Para Berry (2012) esse movimento poderia ser interpretado como uma "terceira onda"<sup>8</sup> das HDs, na qual haveria para ele um "caminho experimental concentrado em torno da computacionalidade subjacente das formas mantidas dentro de um meio computacional" (BERRY, 2012, p. 323, tradução nossa). Ou seja o autor identifica que a reflexão sobre o campo proporcionará um novo nicho que se prestará a um incremento epistêmico, assim para ele o "componente digital das humanidades digitais à luz de sua especificidade mediana", como forma de pensar como as "mudanças mediais produzem mudanças epistêmicas" (BERRY, 2012, p. 323, tradução nossa). Na nossa visão essa nova onda serve de justificativa para uma problematização não superficial do campo, que não se furte a considerar compreender mais sobre seu domínio e escopo.

---

<sup>8</sup> Para Berry (2012) "as *digital humanities* se enquadram em momentos distintos, considerando nessa visão o momento das *Humanities Computing* como sendo a "primeira onda" ocorrida no campo. Segundo ele a primeira onda de trabalho em Humanidades Digitais foi quantitativa, mobilizando os poderes de busca e recuperação do banco de dados, automatizando a linguística de corpus, empilhando *hypercards* em *arrays* críticos. A segunda onda é qualitativa, interpretativa, experiencial, emotiva, geradora de caráter. Na qual se aproveita os kits de ferramentas digitais a serviço das principais forças metodológicas das Humanidades: atenção à complexidade, especificidade do meio, contexto histórico, profundidade analítica, crítica e interpretação. Presner argumenta ainda que a primeira onda de estudos em Humanidades Digitais no final da década de 1990 e início da década de 2000 tendeu a se concentrar em projetos de digitalização em larga escala e no estabelecimento de infraestrutura tecnológica. Enquanto a atual segunda onda de Humanidades Digitais - o que pode ser chamado de "Humanidades Digitais 2.0" - é profundamente generalista, criando os ambientes e ferramentas para produzir, curar e interagir com o conhecimento *born-digital* que vive em vários contextos digitais. A primeira onda concentrou-se, talvez um pouco, na análise de texto (como sistemas de classificação, marcação, codificação de texto e edição colaborativa) dentro das disciplinas estabelecidas. A "*Digital Humanities 2.0*" introduz paradigmas disciplinares inteiramente novos, campos convergentes, metodologias híbridas e até mesmo novos modelos de publicação que muitas vezes não derivam ou limitam-se à cultura de impressão". (BERRY 2012, não paginado, tradução nossa)

Essa nova onda defendida por Berry (2012) pode ser cotejada com a afirmação de SCHREIBMAN; SIEMENS; UNSWORTH (2004, não paginado, tradução nossa) de que a investigação crítica envolve a aplicação de processos de busca, recuperação e crítica, facilitados por algoritmos, que, originados em trabalhos baseados em humanidades, demonstraram ter uma aplicação muito além. Outrossim,

Associada à teoria crítica, essa área é tipificada por estudos interpretativos que auxiliam na nossa compreensão intelectual e estética dos trabalhos humanísticos. Envolve também a aplicação (e aplicabilidade) de ferramentas críticas, interpretativas e algoritmos analíticos, como discutido por Bradley sobre os artefatos produzidos através de processos associados à representação arquivística, disponibilizados por meio de recursos associados a processos de publicação e à comunicação de resultados. Manifestadas nas técnicas de análise discutidas por Burrows e Ide - e vendo utilidade em uma ampla variedade de aplicações, desde a atribuição de autoria (Craig) à estilística cognitiva (Lancashire) - a base de tal análise é o corpus codificado e digitalmente armazenado. por estratégias de representação do conhecimento que, em si, são capazes de possuir o que poderíamos chamar de "poética" (Winder). O mesmo ocorre com as mídias digitais, como o cinema (Kolker), e com as questões de interface e usabilidade que, como discute Kirschenbaum, integram todos os materiais em formato eletrônico e nossa interação com eles. Além disso, os esforços de disseminação têm suas raízes, em última análise, em questões relacionadas à rerepresentação, mas são manifestados em preocupações pertinentes à natureza das comunidades facilitadas por computador (Willett): a preservação em meio eletrônico (discutido por Abby Smith), a publicação eletrônica (tratada pelo capítulo de Jensen, e abordada ainda mais por Palmer), e a matriz única de desafios e oportunidades que surgem com a disseminação das bibliotecas digitais, conforme descrito por Besser. (SCHREIBMAN; SIEMENS; UNSWORTH 2004, não paginado, tradução nossa).

Dessa forma entendemos que uma teorização crítica do arcabouço teórico-conceitual das HDs é altamente pertinente por possibilitar uma robusta problematização das causas e efeitos provenientes das transformações sofridas nesse contexto. Esse é um processo que vai ao encontro da tese da terceira onda de Berry (2012) que potencializa o campo na compreensão dele mesmo, sobretudo em relação às transformações peculiares nas disciplinas das humanidades. É nesse espaço que encontramos oportunidade para discutir sua característica interdisciplinar que pode prover avaliações peculiares para cada área impactada pelas humanidades digitais.

#### **2.4 Arcabouço prático das humanidades digitais: uma parte da investida para compreensão do conceito**

O esforço de propor um diagnóstico do campo das HDs e seu contexto pode não lograr o devido êxito se considerarmos a complexa epistême de construção desse

conceito, além de sua incipiente maturação conceitual. Dessa forma propomos voltar nossa análise para o conjunto de práticas que vem se desenvolvendo no âmbito deste domínio a fim de compor um segundo nível de identificação do que pode ser compreendido enquanto *digital humanities*. Nesse sentido nos associamos à tese de BOREK et al (2016, p.1, tradução nossa) que "classificar e categorizar as atividades que compõem as HDs tem sido uma área de interesse de longa data" para muitos profissionais nesse campo, alimentada por "tentativas contínuas de definir o domínio tanto dentro da academia quanto na esfera pública". Assim, ao nosso ver, está no arcabouço prático também uma fundamental estratégia de identificar o que se abarca enquanto humanidades digitais.

A principal estratégia de identificação do conjunto de práticas que constituem esse campo deve se recair sobre a construção de taxonomias referenciais ao universo das HDs. Nesse sentido o surgimento de diretórios que cruzam fronteiras disciplinares tradicionais também "estimulou o interesse na categorização, com o objetivo prático de ajudar os pesquisadores a identificar, por exemplo, projetos que adotam abordagens, mesmo que o assunto seja muito diferente" (BOREK et al, 2016, p.1, tradução nossa). Dessa forma o autor entende que construção de taxonomias, além de colaborar para compreensão do campo, "também cumprem, na nossa visão, a função de defini-lo", assim para ele, a taxonomia dos métodos de pesquisa digital descritos baseia-se no trabalho anterior em direção a uma abordagem estruturada e baseada em princípios de visão geral do complexo campo das HDs (BOREK et al, 2016, p.2, tradução nossa).

O objetivo de fornecer uma orientação e um meio de pensar o campo estão no centro da ideia de McCarty e Short de "commons" metodológicos. Usando a metáfora e a ferramenta de "mapeamento" para representar o complexo "terreno" do humanidades digitais, McCarty & Short (2002) sugerem um mapa que tem o "commons" metodológico no seu centro. Eles definem os "commons" metodológicos como "uma abstração para os métodos computacionais que as várias disciplinas de aplicação compartilham", funcionando como um espaço de encontro entre "grupos disciplinares" e "áreas de aprendizagem". Os métodos de pesquisa não são nomeados, mas são representados por meio de amplos tipos de dados com os quais estão associados, como "texto narrativo", "imagens" ou "música" (BOREK et al, 2016, p.2, tradução nossa).

A argumentação pela existência de um "commons" fortalece a tese de que um campo que reúne diversos indivíduos de diversas origens epistemológicas necessita realizar um esforço para transformar esse ativo em mais um "commons". Dessa forma a abstração para os métodos computacionais que as várias disciplinas de aplicação compartilham precisam estar expostos para se configurarem um valor compartilhável.

Assim, segundo Borek et al (2016), com um objetivo de estruturação e construção de uma ideia abstrata de empreendimentos de pesquisa individuais identificado em Unsworth (2000), por meio da ideia de "*Scholarly Primitives*", que pode ser considerado atividades básicas para o conhecimento em todas as áreas, foi utilizado como base para a concepção da TaDiRAH enquanto atividades fundamentais que formam “funções básicas comuns aos acadêmicos e atividades através das disciplinas, ao longo do tempo, e independente da orientação teórica” (BOREK et al 2016, p.2, tradução nossa).

Nesse contexto vimos emergir a TaDiRAH - *Taxonomy of Digital Research Activities in the Humanities* como instrumento fluorescente do que Borek et al (2014) considera estar apropriado à compreensão das *digital humanities* sendo melhor entendido através da interseção de seu objeto, metodologias, e aplicativos. Definida pelos autores (2014, p.1, tradução nossa) como "uma taxonomia compartilhada da pesquisa em humanidades digitais", na qual se pode identificar "objetivos e métodos" (por exemplo, captura, análise), "objetos" (por exemplo, dados, imagens, manuscritos), e "técnicas" (por exemplo, análise de cluster, codificação, modelagem de tópico) criadas com a finalidade de colmatar o fosso entre os atores das HDs e suas práticas relacionadas.

À vista disso, podemos considerar, enquanto os três elementos estruturais das HDs, que todo o conjunto prático-teórico deve partir deste eixo fundamental, pois configuram a estruturação, organização e disposição segundo uma ordem ou mesmo a construção embasada por tais orientações, provenientes do esforço deste instrumento taxonômico.

No processo de desenvolvimento da TaDiRAH, extraímos três fontes principais: 1) o artshumanities.net taxonomia para projetos, ferramentas, centros e outros recursos do DH, especialmente expandido por digital.humanities@oxford no Reino Unido e DRAPler (Digital Research and Projects na Irlanda); 2) as categorias de DiRT para ferramentas de pesquisa digital, relançadas sob o Projeto Bamboo nos EUA, que segue contínua mesmo depois do final do projeto; e 3) o esquema usado pela bibliografia do DARIAH "Doing Digital Humanities" estruturada via Zotero para organizar a literatura sobre todas as facetas das humanidades digitais. Esses recursos foram mapeados, analisados e destilados em suas partes essenciais, produzindo uma taxonomia simplificada de dois níveis: oito “objetivos” de alto nível que são amplamente baseados no processo de pesquisa acadêmica e vários “métodos” de nível inferior associados a cada meta. Além disso, existem duas listas abertas separadas de “objetos” de pesquisa em humanidades digitais e “Técnicas” que podem ser livremente associadas a métodos de alto nível. (BOREK et al, 2014, p.181-182, tradução nossa).

Por certo, podemos assumir que a TaDiRAH corresponde à classificação mais expressiva e fundamentada sobre as atividades de pesquisa digital nas humanidades, certificando senão todos, a parte mais significativa da estrutura das humanidades digitais,

se qualificando como principal expoente taxonômico para área o que justifica sua abordagem enquanto recurso complementar à compreensão deste domínio. Dessa forma, embasados por ela, entendemos as HDs enquanto composição da seguinte estrutura:



Figura 5: Diagrama proposto para representação material das *digital humanities*. (Fonte: elaborado pelo autor baseado na TaDiRAH).

Dessa maneira alcançamos, ao nosso ver, enquanto sumário deste domínio, a estrutura empírica central do arcabouço prático-conceitual das HDs representado pela tríade: atividades de pesquisa, técnicas de pesquisa e objetos de pesquisa. Compondo as atividades de pesquisas, compreende-se as ações que caracterizam as práticas do campo que, amparadas pelas técnicas, ou seja, os procedimentos que irrompem como produto das ações/atividades mediadas pelas tecnologias, exponenciam o novo conceito. Por fim, essas funções baseadas necessariamente no objeto de pesquisa, exclusivamente enriquecido digitalmente de uma ontologia conferida pela ontologia do digital.

De certo este trinômio estrutural é a harmonização mais objetiva que podemos alcançar. Por ser um instrumento construído colaborativamente em um domínio multidisciplinar, a TaDiRAH enfrenta desafios para seguir representando as práticas em HDs, como aponta BOREK et al (2014, p.182, tradução nossa) "a preocupação em evitar a confluência de conceitos, conciliando termos, minimizando a redundância, equilibrando a “correção” teórica contra a necessidade de adotar termos comumente usados para garantir uma localização a outro"

Uma vez alcançada uma estrutura empírica abrangente, podemos dispor de subsídios para compreensão mais acurada do campo, impedindo que o entendimento multidisciplinar predominante da área seja responsável por tolir qualquer iniciativa de identificação das características desse domínio. Apesar de uma análise mais acurada, voltada formalmente ao domínio, essa estrutura empírica (fig 5) permite apontar para o horizonte que pretendemos. Como temos, assertivamente, a ideia da multidisciplinaridade

como pedra angular do gene das HDs, todavia, apesar dessa concepção, não podemos naturalizar que todos os agentes desse processo possuam uma equidistância. Por equidistância, nos baseamos na ideia de participação no movimento das *digital humanities* em que conjuguem uma maior proximidade com o conjunto de elementos trabalhados pelo domínio. Ou seja, tomando por base a estrutura material desse movimento, construímos condições menos abstratas de perceber, no que se cunha como HDs, a nítida marca de outras disciplinas.

Nesse sentido, ao vislumbrar o que vem a se considerar HDs em termos usuais, podemos inferir que determinados campos e/ou ciências podem estar espelhados nesse domínio. Como pontuam VANDEGRIFT; VARNER (2013, p.2, tradução nossa) tão diferentes quanto esses usos podem ser, um "'fio condutor' comum que passa por cada um deles é a percepção de que a tecnologia está permitindo que o trabalho das humanidades seja mais envolvente e mais acessível". Todavia a experiência de cada uma delas nesse "fio condutor" estará ligada intimamente a proporção de seus objetos originais de estudo. Assim, por exemplo, os autores destacam que bibliotecas estão passando por mudanças revolucionárias à medida que a tecnologia da informação evolui possibilitando que "a biblioteca pode reinventar o seu lugar no ciclo e produção acadêmica" devido a um "emparelhamento com o impulso do digital nas humanidades" (VANDEGRIFT; VARNER, 2013, p.2, tradução nossa).

Dessa forma nos amparamos na proposição de análise do elemento comum da estrutura empírica diagramada pela figura 5 '*research objects*' como parte determinante do que defendemos: de que não há uma equidade entre os "membros" das humanidades que compõem atualmente as HDs. Além de estarem formalmente cunhados pela TaDiRAH como objetos de pesquisas das *digital humanities* os mesmos estão elegíveis, desde outrora, enquanto elementos de pesquisa das bibliotecas e unidades de informação em geral, concomitantemente às áreas da informacionais. Spiro (2011) *apud* VANDEGRIFT; VARNER (p. 3, tradução nossa) centra-se em cinco áreas onde HDs focam, além da atividade acadêmica tradicional: "fornecimento de amplo acesso à informação cultural; habilitação e a manipulação de dados; transformação da comunicação acadêmica; aprimoramento do ensino e a aprendizagem; e promoção de impacto público". Assim nos é possível inferir que o papel das ciências da informação nas HDs é indissociável, pois tanto seus objetos de estudo quanto a heurística aplicadas a estes estão no foco das *digital humanities* por se constituírem '*research objects*' da representação empírica deste domínio.

De certo vislumbramos espaço para advogar a construção de uma heurística comum e íntima, diferente da relação dos demais membros desse domínio. Como identificam os autores

Para aqueles de nós que vivem em bibliotecas, esses objetivos devem soar muito familiares como a preocupação com o acesso e impacto público que estão entre nossos valores mais importantes. Essa convergência de valores provavelmente explica por que já existe uma rica tradição de parcerias bibliotecárias em projetos de humanidades digitais. De fato, alguns dos projetos mais antigos e mais venerados são essencialmente arquivos on-line e têm sido produzido por parcerias entre docentes e bibliotecas. [...] A emergência de recursos no âmbito das bibliotecas como repositórios digitais, projetos de dados cooperativos vinculados, habilitados tecnologicamente e ambientes de trabalho colaborativos, oferecem novas maneiras para os acadêmicos trabalharem com bibliotecas e não apenas neles. (VANDEGRIFT; VARNER, 2013, p.3, tradução nossa).

Por essa razão sinalizamos que a heurística, já concretizada como humanidades digitais, pavimenta uma nova construção derivada da detecção de uma relação peculiar, porém muito particular, entre os membros de sua interdisciplinaridade. Assim, essa relação plural deve ser explorada à medida que as reflexões acerca das HDs se desenvolvam.

## **2.5 As humanidades digitais sob evolução epistêmica de sua própria heurística**

Podemos constatar na pluralidade das *digital humanities* e em sua característica interdisciplinar um facilitador para um acercamento entre áreas, temas e objetos. Como pregam Tang; Cheng; Chen (2017) as HDs apresentam um terreno fértil e apropriado para o estudo da integração do conhecimento por focar diretamente nos problemas pragmáticos de como a computação auxilia na prática de pesquisa e como problemas teóricos e práticos recebem nova perspectiva através da computação. Essa nova perspectiva demanda problematizações, pois alheio ao *statement* de construção de um domínio genuinamente novo, podemos relativizar os movimentos das HDs enquanto uma narrativa autoafirmativa que pode tender a extenuar-se. Esse cenário pessimista pode ser potencializado basicamente por dois motivos. O primeiro teria relação com a ausência de um aprofundamento epistemológico sobre o movimento das *digital humanities*, enquanto o segundo estaria ligado a uma reflexão no sentido de que, passado os instantes áureos de reverberação da marca '*digital humanities*', as humanidades não acabariam por se dar conta que, uma vez rompido esse debate, não haveria mais espaço para uma tradição



metodológica que não englobasse compulsoriamente a tecnologia como instrumento fundamental. Todavia esses dois motivos possuem relação reciprocamente causal.

Partindo do cenário mais pessimista, é inevitável conjecturar que esse é um processo que tende a ganhar fôlego. Convencidos de que tanto a matéria-prima de pesquisa quanto todo o ferramental à serviço do ensino, pesquisa, educação, produção de conhecimento, etc migra em escalas exponenciais para o computadorizado, na qual a predominância dos computadores e softwares na estrutura acadêmica das humanidades pode levar à percepção de que as HDs não correspondam mais a um movimento independente, mas sim a uma compulsoriedade. Logo o raciocínio coletivo dos humanistas não interpretaria mais o movimento com distinção, mas sim como uma simbiose.

Essa simbiose parece, ao nosso ver, já ter ocorrido no âmbito de disciplinas que sofreram esse tipo de fenômeno a longo prazo e em outras proporções. É o que VANDEGRIFT; VARNER (2013, p. 4, tradução nossa) consegue ilustrar ao destacar que "à medida que a biblioteca continua a evoluir, ela deve funcionar cada vez mais como um lugar onde os estudiosos podem experimentar novas coisas, explorar novas metodologias e geralmente experimentar novas maneiras de fazer estudos". Por isso a ideia de simbiose parece, por exemplo, estar mais acurada com as disciplinas da Ciência da Informação, pois a natureza do campo e seu objeto já vem sendo preparado enquanto laboratório para o que se versa como humanidades digitais. Os autores completam esse argumento quando pontuam que como recursos se movem para o digital e o espaço é realocado de pilhas para "laboratórios" e "objetos transversais", o argumento de que a interação com esse cenário não é imperativa para as tarefas críticas da biblioteca moderna. (VANDEGRIFT; VARNER 2013, p. 4-5, tradução nossa).

Por hora é neste contexto que nos detemos, numa simbiose que possibilita identificar o movimento das HDs como parte intrínseca das humanidades enquanto sinal do nosso tempo. Todavia, essa conjectura necessitará de um amadurecimento epistêmico que exigirá um aprofundamento nas reflexões sobre o campo. A princípio a Ciência da Informação mais uma vez se credencia a contribuir também nessa proposta. Como os estudos desse movimento têm, no campo, total extemporaneidade de *expertise*, homologam-se a acrescentar às *digital humanities* ou aos humanistas digitais. Por ser um dos pontos nevrálgicos das HDs a delegação ao autômato da atribuição de tarefas, implica-se numa intervenção na relação indivíduo e informação. A essa intervenção

entendemos como instituição de uma mediação que, desde outrora, é trabalhada no âmbito da Ciência da Informação num contexto altamente respectivo.

Por fim, julgamos que essas indicações, além de contribuírem para consolidação do campo das humanidades digitais, revelam uma relação própria que pode representar através da complexa ideia de interdisciplinaridade, um fator que nos leva a cada vez mais argumentar que as áreas em análise produzem uma justaposição. Por isso esse estudo se empenhou no esforço de topografar as humanidades digitais a partir de sua literatura, para alcançar alguma mensuração da presença e da *expertise* da Ciência da Informação em sua natureza.

### **3 UM FOCO NA ESTRUTURA ELITISTA DA CIÊNCIA PARA COMPREENSÃO DESTE DOMÍNIO: UMA RESSONÂNCIA DOS ESTUDOS BIBLIOMÉTRICOS**

Os esforços para compreender o campo das humanidades digitais são complexos, a recente formação da área sob sua alcunha oficial e a marcante característica de uma profunda composição interdisciplinar, conferem à esta tratativa uma complexidade ímpar. Após centrar esforços numa identificação conceitual para a área, que tem em sua ausência de consolidação epistêmica mais um fator de complexidade, empreendemos uma sumarização deste domínio através de uma estrutura empírica central, à luz de seu arcabouço prático-conceitual, baseado em seu principal instrumento taxonômico. Destarte consideramos analisar outros indicativos das *digital humanities* como forma a compreender em maior profundidade seu universo.

Muito embora a ação seja altamente passível de uma inclusão, dispomos de métodos historicamente difundidos, sobretudo no universo da Ciência da Informação, para empreender uma análise com essa proposta. Todavia a ação seja complexa e seu resultado altamente passível de incompletudes, os resultados dessa iniciativa nos conferem um caminho que, por mais que seja volúvel, nos subsidia na construção de alguns indicativos que embasem a ideia central encampada proposta por essa pesquisa.

Dessa maneira identificamos nas técnicas das análises bibliométricas maneiras para compreender o campo via a produção de seus agentes, que gerando conhecimento e reflexões sobre o mesmo, influenciam em seu desenvolvimento, imprimindo nele suas características e vieses. Esse esforço resulta inevitavelmente em algum tipo de compreensão do domínio, nesse caso no das HDs. Nesse sentido essa foi a estratégia julgada apropriada enquanto metodologia para a pesquisa nesta tese apresentada.

Para atingir tamanho objetivo, este capítulo se presta a jogar luz sob o fator elitista da ciência, que a despeito de fundadas análises sobre tal universo que exploraremos neste capítulo, nos garante a efetividade de, ao investigar esse domínio, estar em garantia de fazê-lo a partir da busca, identificação e análise de tal "núcleo eloquente" de produtores de conhecimento.

Dessa forma a clarificação do campo das HDs deve se dar por meio da imputação de uma análise básica do domínio, produzindo subsídios a partir do estudo bibliométrico, para identificar atores de maior eloquência, influência e relevância destes ou de recursos

por eles produzidos que dão corpo e contorno ao campo. Assim sendo advogamos a construção de um diagnóstico do campo das humanidades digitais por tais vias.

### **3.1 Uma abordagem bibliométrica como caminho para um diagnóstico das humanidades digitais**

O movimento pelo qual os mais diversos campos científicos sofrem seu processo de consolidação epistêmica passa necessariamente pela atividade da comunicação científica. É neste contexto que o desenvolvimento de novos conhecimentos iniciam seu processo de avaliação pelos pares, saindo de estudos teóricos individuais para creditação coletiva destes. Como CARIBÉ (2015, p. 90) pontua "a comunicação científica integra o processo de produção e desenvolvimento da ciência, tão vital como a fase de coleta e análise dos dados". A ciência, pela sua própria natureza, representa o conjunto de conhecimentos públicos e, ao nosso ver, o processo da comunicação na ciência se caracteriza enquanto processo vital que desenvolve o conhecimento a partir de elos construídos entre os pares, imputando uma impressão de influências que se tornarão marcas daquele determinado campo. Assim, para a autora, este processo pode ser entendido como "qualquer atividade ou comportamento que facilita a construção e o compartilhamento de significados entre indivíduos [...] consistindo no conjunto de relacionamentos entre os indivíduos unidos pelos significados que constroem e compartilham entre si" (CARIBÉ, 2015, p. 90).

Esse movimento se trata de um processo complexo que desdobra sua abrangência em 3 fases, segundo Lievrouw; Carley (1990) *apud* Caribé (2015), correspondendo à concepção, documentação e popularização. Em seu primeiro estágio, a autora identifica que existe "grande quantidade de informação tanto científica quanto social; interesses, conceitos e métodos; um único paradigma científico. Utilizam discurso comum [...] as estruturas de comunicação são informais, o que facilita a interação interpessoal" (CARIBÉ, 2015, p.91). Por se tratar de um estágio primário vemos que a característica marcante dessa etapa é a interação pessoal. A relação entre indivíduos é marcante e corrobora nossa visão de que essas marcas projetam-se no desenvolvimento do campo tornando-o, naturalmente, assemelhado com os que o integram. Todavia, essa característica se imprime, predominantemente, nesta etapa inicial, pois na etapa de documentação, se imputa uma profissionalização dos saberes e as vias científicas passam

a ditar o protocolo para expansão do conhecimento. De certo a autora identifica que no estágio da documentação

compartilham muita informação científica, porém há pouca interação social. Partilham conhecimento, métodos e discursos; podem ou não concordar com o mesmo paradigma científico, pois geralmente utilizam múltiplos paradigmas; os canais formais de comunicação são mais utilizados que os informais; os contatos são mais formais e menos frequentes; os grupos são mais heterogêneos, social e culturalmente; e o vocabulário é menos compartilhado. **A natureza formal das estruturas de comunicação (sociedades profissionais) e dos processos (publicar, especialmente em periódicos de interesse) durante a etapa de documentação tende a direcionar a difusão de novas informações. O processo de comunicação é mais formal e produz registros na forma de documentos provenientes de um corpo coerente de pesquisa.** Há certa limitação para a comunicação de ideias, pois o cientista adapta-se às convenções exigidas pelo meio de comunicação ou outros canais utilizados nessa fase. (CARIBE, 2015, p. 91, grifo nosso)

Enquanto foco de nossa abordagem, julgamos que reside nesta etapa - da documentação- a mais relevante das três por ser nesta onde o processo de consolidação se exerce via padrões creditados cientificamente. Assim entendemos que nesta etapa ocorre a mais relevante filtragem em relação a uma outra etapa -no caso a primeira, já que na terceira as ideias científicas podem tornar-se parte do discurso público por meio de diferenciados canais, segundo Caribé (2015). Ou seja, a terceira etapa limita-se a reproduzir uma consolidação executada no âmbito da segunda.

Essas observações lastreiam nossa escolha da análise da literatura enquanto indicativo mais apurado de representação do campo das humanidades digitais. Como identificado nos padrões da comunicação científica, nessa literatura estão impressos, por certo, características peculiares que apenas podem ser desveladas a partir da análise desses registros que se espera constituírem sinais expoentes do campo. Como defendido essa estratégia agrega à compreensão das *digital humanities* não apenas por sua precocidade epistemológica, mas também pela imposição de uma visão supérflua do campo tipo como meramente interdisciplinar. Assim nossa estratégia de realizar uma abordagem bibliométrica se apresenta como aquela com maior retorno tangível a ser empregada neste caso.

Por essa razão a pesquisa centra esforços na busca de um mapeamento da literatura por ser esta um sinal sensível do campo científico e no caso particular das humanidades digitais seu maior substrato representativo. Em igual proporção concordamos que

na medida em que a ciência vai se estabelecendo ao longo do tempo, ressaltando ora um aspecto, ora uma metodologia, ora determinado referencial teórico, o monitoramento dessa produção precisa ser paralelamente elaborado, pois, ao identificar e explicitar os caminhos que a ciência percorreu, torna-se

possível revelar o processo de construção do conhecimento sobre determinado tema, ou seja, identificam-se duplicações, contradições e, especialmente, lacunas, isto é, aspectos ainda não explorados. (CASSUNDE; BARBOSA; MENDONCA, 2018, p. 312)

Por certo, nesse caso, estamos diante de uma complexidade ainda maior, pois além de a ciência sofrer seu processo natural de acercamento por suas contradições e lacunas, o mapeamento da produção das HDs pode prestar um relevante serviço não apenas ao campo, mas também a todos que nele se identificam como parte. Assim, hipoteticamente, com a característica aglutinadora das HDs em torno de uma problemática "nova" ter sido a responsável por conceber esse campo, é conveniente especular se esse é um movimento que tende a uma consolidação em torno de sua configuração atual ou se as áreas podem voltar-se a seus interiores, a partir de então, incorporando tais questões. Por outras palavras é de se pontuar que as humanidades podem passar por um processo de simbiose com aquilo que hoje lhe é considerado alheio e responsável por sua formação enquanto *digital humanities*. Nosso ponto é que será a literatura científica a primeira a denunciar esse eventual movimento e tantos outros sequer especuláveis.

Nesse contexto uma das ferramentas metodológicas mais difundidas para a mensuração da produção científica são os estudos bibliométricos. Essa necessidade se desenvolve, inicialmente, para atender uma demanda quantitativa fruto de exigências de produtividade acadêmica. Como observa ARAÚJO (2006, p.12) esses estudos surgem no início do século "como sintoma da necessidade do estudo e da avaliação das atividades de produção e comunicação científica desenvolvendo-se inicialmente a partir da elaboração de leis empíricas sobre o comportamento da literatura".

Por outro lado, ao se prestar a análises qualitativas, pode se configurar enquanto instrumento essencial para a mensuração de uma comunidade científica. Para CAFÉ; BRÄSCHER (2008, p.54) a Bibliometria pode ser definida como "um conjunto de leis e princípios aplicados a métodos estatísticos e matemáticos que visam o mapeamento da produtividade científica de periódicos, autores e representação da informação". Na mesma medida PRICE (1976, p.39) resume, objetivamente, a essência das práticas bibliométricas por parecer clara a importância de se "dispor de uma distribuição que nos informe sobre o número de autores, trabalhos, países ou revistas que existem em cada categoria de produtividade, utilidade ou o que mais desejarmos saber". Ao nosso ver, essas visões corroboram o caráter fundamental deste tipo de estudo para qualquer área da ciência.

A relevância e abrangência dos estudos dessa natureza remontam a consolidação de sua inserção no contexto da Ciência da Informação, quando o termo foi inicialmente cunhado, enquanto descrição de todos os estudos que buscam "quantificar o processo da comunicação escrita na ciência" (PRITCHARD *apud* RAO, 1986, p. 179) ou como o "tratamento quantitativo das propriedades do discurso registrado e do comportamento referente ao mesmo" (FAIRTHORNE *apud* RAO, 1986, p. 179). Assim, para o autor, a expressão mais representativa da bibliometria abarcava a ideia de uma abordagem dos processos de informação e seu tratamento, utilizando-se da análise quantitativa das características e do comportamento dos documentos.

Para Rao (1986) na base das abordagens bibliométricas se encontram as distribuições usadas para tratar tais estudos. Dessa forma estas distribuições se estruturam na "frequência da ocorrência de palavras em texto; na produtividade de autores em termos de documentos científicos e além de na dispersão dos artigos em diferentes publicações periódicas" (RAO, 1986, p. 180). À vista disso, Café; Bräscher (2008) apresentam 3 grandes balizadores empíricos, que para nossa abordagem se aplicam enquanto institutos teóricos, considerados expoentes dos estudos bibliométricos e abrangem as principais unidades de análises praticados nessa esfera. Nesse contexto nos baseamos nas Leis de Bradford, Lotka e Zipf que teorizam o incremento da produção científica e seus impactos no âmbito dos periódicos, autores e palavras respectivamente. Essas leis,

sintetizam-se da seguinte forma: Bradford objetiva conhecer o núcleo de periódicos produzido em determinado tema, Lotka visa definir as maiores contribuições de pesquisadores em determinadas áreas do conhecimento e Zipf pontua a frequência com que certas palavras aparecem nos textos científicos de maneira a definir sua representatividade neste contexto. (CAFÉ; BRÄSCHER, 2008, p. 54)

Assim temos na bibliometria um robusto método de mapeamento e análise da literatura que preza, grosso modo, três pontos centrais: a compreensão das atividades de um campo em suas estruturas alicerçantes como na produção científica de seus agentes (publicações); nos meios de ventilação dessa produção (periódicos); e, não obstante, na estratégia de mensuração do significado desse volume produzido expressado pela recorrência de suas palavras (palavras-chave). Decerto a definição desse escopo nos sustenta enquanto método para produzir uma análise do *corpus* da literatura das *digital humanities* como expressão de suas características.

Dessa forma vemos fundamentação de nossa metodologia de identificação de termos dessa literatura na Lei de Zipf que segundo Rao (1986) estabelece relação entre a

posição de uma palavra e a frequência de seu aparecimento em um texto longo. O autor explica que Zipf extraiu sua lei de um princípio geral do esforço mínimo, pontuando que "palavras cujo custo de utilização seja pequeno ou cuja transmissão demande um esforço mínimo são frequentemente usadas em um texto" (RAO, 1986, p. 180). O autor explica que sua formulação encontra expressão na equação  $rf=c$ , na qual 'c' é uma constante, 'r' é a posição da palavra e 'f' sua frequência. Assim, temos na 'Curva de Zipf', a representação gráfica do fenômeno (fig.6)

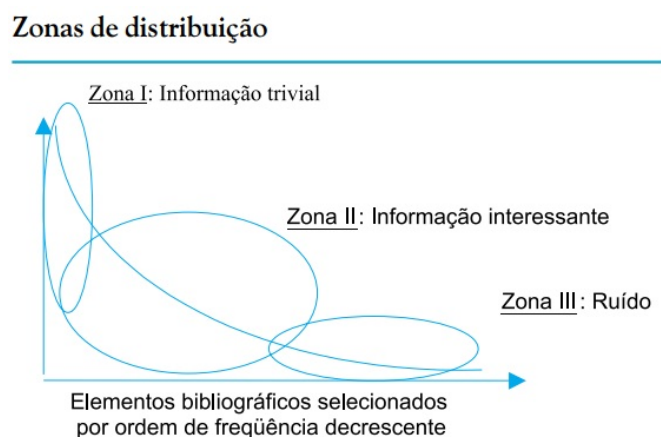


Figura 6: Zonas de Distribuição da Curva de Zipf de acordo com Quoniam (1992) Fonte: Quoniam (2001).

O fenômeno representado graficamente pela Curva de Zipf (fig.6) é descrito segundo QUONIAM (2001, p. 24) pela "setorização de palavras de baixa frequência que ocorrem de modo que várias palavras acabam por apresentar a mesma frequência". De acordo com o autor, a Zona 1 detém a informação trivial na qual se concentram os temas centrais da análise, enquanto que a Zona 2 deterá temas com certa relevância uma vez que contem informação inovadora, mas também temas periféricos. Já a Zona 3 é caracterizada pelo autor como ruído, constituindo-se de termos de pouca relevância, não sendo possível identificar se terão relevância futura. Nesse sentido nossa análise buscará a identificação desses setores de informação presentes no esquema de Zipf, porém baseando-se e verificando a aplicação desta formulação à análise produzida pela ferramenta utilizada para alcançar tal resultado (VOSviewer).

No tocante à produtividade dos autores encontramos argumentação teórica na Lei de Lotka que vai ao encontro das abordagens sobre produtividade científica. Na concepção de sua lei, Lotka, segundo Rao (1986), baseou-se numa análise do número de publicações relacionadas no *Chemical Abstracts*, de 1907 a 1916, com a frequência de



publicações por autores isoladamente. Pareou essa análise com corpus semelhante selecionado no *Auerbach's Geschichtstafel der Physik* e calculou as frequências teóricas de publicações de autores usando o método de mínimos quadrados. A partir desse estudo, Lotka propôs a lei dos quadrados inversos na qual relaciona a "frequência relativa de autores publicando um determinado número de trabalhos". Sendo assim, para RAO (1986, p. 182) a equação de Lotka "é determinada em sua forma geral por três parâmetros: número de cientistas com produtividade mínima (único trabalho); produtividade máxima de um cientista e o expoente característico".

A Lei de Lotka baseia os debates acerca da produtividade científica que, em linhas gerais, é medida frequentemente em termos de produto publicado (Rao, 1986). Algumas abordagens enfatizam relevância apenas no número de publicações de um indivíduo, outras na relação entre quantidade de obras publicadas face seu reconhecimento científico. Esses são alguns exemplos do porquê a Lei de Lotka deve fundamentar as análises a respeito da produtividade científica e seus atores, mas não estar aplicadas a todas elas. Rao (1986) alerta que muito embora ela não seja a única, a Lei de Lotka configura-se como a mais difundida em contextos bibliométricos.

Lotka fornece a base teórica fundamental para estudos de bibliometria que têm a ver com autoria. Lotka estava interessado em determinar "a parte que, homens de calibre diferente, contribuem para o progresso da ciência" [...] Com base no índice Lotka, é costume distribuir os autores de um conjunto específico de publicações em três níveis de produtividade: pequenos produtores (com apenas um emprego e índice de produtividade igual a 0), médios produtores (entre 2 e 9 postos de trabalho e índice de produtividade maior que 0 e menor que 1) e grandes produtores (10 ou mais empregos e índice de produtividade igual ou maior que 1). (SPINAK, 1996, p. 147-148, tradução nossa).

De toda forma, a base da lógica de Lotka se inspira, de acordo com Rao (1986), na sugestão de que o aumento  $x$  é acompanhado do aumento de cientistas de baixa produtividade e que estes devem corresponder à maioria do corpo produtor de conhecimento do campo. Assim, para ALVARADO (2002, p. 14), "o número de autores que fazem  $n$  contribuições em um determinado campo científico é aproximadamente  $1/n^2$  daqueles que fazem uma só contribuição e que a proporção daqueles que fazem uma única contribuição é de mais ou menos 60%". Dessa maneira, mais uma vez, nosso objetivo é constatar a aplicação deste princípio à análise deste mesmo espectro dos autores à análise possibilitada pela ferramenta utilizada por esta pesquisa. Assim pretendemos identificar via VOSviewer o grupo produtor responsável pela maior produtividade do campo considerando as variações das diversas métricas relacionadas à autoria.

Uma outra nuance da mensuração da produção de literatura relaciona-se com os periódicos científicos. Instrumentos de disseminação de informações, correspondem a um tipo de vetor de conhecimento ao passo que divulgam experimentos e estudos, legitimando determinado saber ao colocá-los ao alcance da comunidade acadêmica. Creditado muitas das vezes por pares, esses estudos projetam seus autores e os conferem reconhecimento acadêmico, credenciando esses indivíduos a uma proeminência no campo. Nesse sentido, encontramos espaço para abordar a Lei de Bradford ou da Dispersão que, em larga medida, reclinam-se sobre a mensuração da produção científica face seu principal vetor formal. De acordo com RAO (1986, p.186) Bradford descreveu um "quadro da dispersão de periódicos na área de geofísica, no qual representou as somas parciais das referências em função do logaritmo natural da soma parcial de número de periódicos". Segundo o autor, Bradford constatou que o gráfico resultante é uma linha reta. Baseado neste experimento, Bradford sugeriu a formulação ' $F(x)=a+b \log x$ ' para descrever um fenômeno de dispersão (RAO, 1986, p. 186) onde,

a lei de distribuição de artigos em um dado assunto em revistas científicas podem ser estabelecidos a partir seguinte maneira: se as revistas científicas estão ordenadas em seqüência decrescente de produtividade de artigos em um determinado tópico, estes podem ser divididos em um núcleo de revistas que se dedicou mais em particular ao assunto e vários grupos ou zonas contendo o mesmo número de itens que o núcleo, onde o número de periódicos no núcleo e as zonas sucessivas estarão na proporção de 1: n: n<sup>2</sup> (Bradford *apud* SPINAK, 1996, p. 38, tradução nossa).

Em sua formulação para a dispersão, Bradford considera 'F(x)' enquanto número acumulativo de referências contidas no periódico 'x' mais produtivo, sendo 'a' e 'b' constantes. A representação gráfica desta formulação pode ser compreendida pela figura (7) abaixo

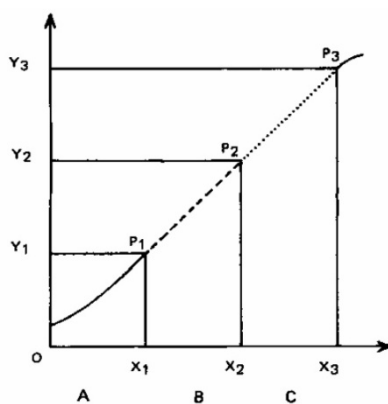


Figura 7: Representação gráfica para formulação da Lei de Bradford. Fonte: Pinheiro (1983).

Para PINHEIRO (1983, p.63) "a parte inicial é correspondente à concentração, restrição de Bradford (zona A); a do meio, de produtividade média, é a componente de Zipf (zona B) e a terceira (zona C), compreende os periódicos de baixa produtividade". Dessa forma podemos identificar similitudes com a abordagem da Lei de Zipf por Bradford também localizar 3 setores, nos quais a concentração de relevância de um tema recairá sobre um grupo estreito de periódicos. Nas palavras de PINHEIRO (1983, p. 62-63), Bradford sinalizou a relevância das zonas de sua lei em "aqueles que produzem mais de quatro referências por ano"; "os que produzem mais de uma e não mais do que quatro referências, anualmente"; e os "periódicos que produzem uma referência ou menos por ano". Assim,

Pelo enunciado da lei: "se os periódicos forem ordenados em ordem de produtividade decrescente de artigos sobre um determinado assunto, poderão ser distribuídos num núcleo de periódicos mais particularmente devotados a esse assunto e em diversos grupos ou zonas contendo o mesmo número de artigos que o núcleo, sempre que o número de periódicos e das zonas sucessivas for igual a  $1:n:n^2$  (PINHEIRO, 1983, p.62-63)

Nesse sentido temos na Lei de Bradford um importante balizador teórico que nos permite corroborar a tese de que a produção das humanidades digitais concentra-se em um grupo seletivo de periódicos e que, ao identificar as áreas das quais estes periódicos são representantes, teremos a garantia da hipótese de que essas áreas são, em larga medida, constituidoras eloquentes desse campo. Assim significa dizer que como nas leis balizadoras da bibliometria que explicitam, basicamente, a ideia de concentração, podemos, através das análises propostas nesta pesquisa, identificar quais áreas formam em maior concentração as *digital humanities* constituindo um tipo de hegemonia ante as demais.

Além dos diversos aspectos ressaltados sobre a relevância dos estudos bibliométricos, destacamos aquela que talvez seja um de seus produtos mais expressivos que é a identificação da concentração da produção da literatura em indivíduos que poderão ser considerados como os expoentes de determinado campo ou área. Corroborando esse ponto de vista, SPINAK (1996, p. 72, tradução nossa), cita a lei de Lotka como exemplo "na qual se vê que poucos autores são prolíficos e a maioria produz apenas um ou duas contribuições no máximo" e ainda que "embora algumas entidades sejam responsáveis pela produção de grande parte dos documentos, muitos deles estão espalhados entre muitas entidades menos prolíficas". Para o autor a bibliometria permite identificar a concentração da produção e sua simultânea dispersão. Nesse sentido focamos

na proposta de desvelar esse microuniverso concentrador da produção das *digital humanities* através da utilização do software VOSviewer.

Várias outras abordagens se dedicam a problematizar a questão da concentração de produção acadêmica exatamente por ser este um fenômeno passível de diversas variáveis. Autores como Spinak (1996) nomeia essa manifestação enquanto uma estrutura elitista da ciência, segundo sua interpretação de conceituações no âmbito social da ciência como, por exemplo, nas Hipóteses de Ortega<sup>9</sup> e Eclesiastes<sup>10</sup>, que em linhas gerais, versam que "a concentração de 80% produção científica relevante em apenas 20% da produção pesquisadores abre margem a uma extensa discussão sobre financiamento na ciência" (SPINAK, 1996, p. 93, tradução nossa). Outros contributos teóricos foram importados para bibliometria no esforço de problematizar esse caráter concentrador da produtividade acadêmica, como o Índice de Gini<sup>11</sup>, que embasaria segundo SPINAK

---

<sup>9</sup> Na percepção do Cole (hipótese de Ortega), um pequeno grupo de "Newtons e Einsteins" (a elite) deve receber o maior parte dos recursos, então, de acordo com os índices de cotação, o resto da comunidade científica é dispensável. A política científica deve ser seletiva e elitista. Eles consideram que seria redução conveniente do número de centros de treinamento e doutorado (Ph.D.) para apenas algumas das universidades de maior prestígio. Preocupado com o mercado de trabalho disponível para a massa de doutorados em ciência sugerem que uma redução de 50% na equipe científica pode resultar em salários, carreira e prestígio científico mais atraentes, sem grandes perdas para o avanço da ciência em geral. (SPINAK, 1996, p. 94, tradução nossa)

<sup>10</sup> Explicação alternativa apresentada por Turner e Chubin (1976, 1979) à hipótese de Ortega apresentada por Cole (1972) em relação à estrutura elitista da ciência. De acordo com a interpretação de Cole, a concentração de 80% da produção científica relevante em apenas 20% da produção pesquisadores apoiariam as seguintes conclusões com as quais políticas científicas devem ser redefinidas. O mesmo talento científico poderia ser usado se fosse usado apenas 20% dos cientistas mais destacados e não recursos de resíduos que suportam os outros 80%, (estrutura elitista da ciência) de acordo com os índices de citação, o sistema de ciência estratificada é "justa". A visão do Cole é baseada em uma explicação funcionalista da ciência, onde se considera que todos os cientistas podem ser substituídos, alguns podem ser substituídos mais rapidamente, já que os outros são menos substituíveis. Em suas palavras: "Embora todos os cientistas sejam substituíveis no sentido que outros cientistas acabariam por duplicar descobertas, alguns cientistas têm mais equivalentes funcionalidades do que outras "Turner e Chubin apresentam uma explicação alternativa, que eles chamam a hipótese de Eclesiastes, derivada das palavras do verso 9:11 daquele livro da Bíblia: *eles não têm a corrida rápida, nem os poderosos a batalha, nem os sábios têm o pão, nem sequer têm o pão riquezas entendidas, nem mesmo aquelas que têm habilidade eles têm o favor; porque o tempo e o imprevisto eles acontecem com todos*. De acordo com essa hipótese, a estratificação emerge como reflexo das estruturas das organizações, dos círculos de influência e eventos incidentais (estar no lugar certo no tempo adequado) do que para diferenças marcantes na capacidade científica entre as pessoas. As diferenças de talento seriam genuínas, porém marginais. De acordo com essa hipótese, o fato de que 80% dos pesquisadores têm apenas 20% do impacto e isso implicaria que os talentos são largamente desperdiçados. (SPINAK, 1996, p. 93, tradução nossa)

<sup>11</sup> Recurso da Economia que pode ser usado para qualquer distribuição embora seja comumente utilizado para medir a desigualdade de distribuição de renda. Segundo o Instituto de Pesquisa Econômica Aplicada (2004) o Índice de Gini, criado pelo matemático italiano Conrado Gini, é um instrumento para medir o grau de concentração de renda em determinado grupo. Ele aponta a diferença entre os rendimentos dos mais pobres e dos mais ricos. Numericamente, varia de zero a um (alguns apresentam de zero a cem). O valor zero representa a situação de igualdade, ou seja, todos têm a mesma renda. O valor um (ou cem) está no

(1996, p. 111, tradução nossa) "uma 'teoria da concentração' ao estudar o grau de variedade de distribuição de recursos, àqueles que podem ser identificados com os conjunto de números". Nesse sentido, a despeito do amplo debate acerca desta problemática, nossa pesquisa reafirma o caráter fundamental da persecução dessa 'estrutura elitista da ciência' no âmbito das humanidades digitais, por residir nela um caráter identitário que, ao nosso ver, fornece paridades concretas estabelecidas com o campo e suas insígnias viscerais.

### 3.2 Ensaando a análise do domínio como produto do diagnóstico bibliométrico

A persecução de uma 'estrutura elitista da ciência' enquanto produto de um esforço amparado metodologicamente por um estudo bibliométrico, nos remete necessariamente ao estabelecimento de uma comunidade. No caso em questão de uma comunidade de discurso. Na visão de NASCIMENTO; MARTELETO (2004, [p.5]) "as comunidades discursivas são distintos grupos sociais sincronizados em pensamento, linguagem e conhecimento, constituintes da sociedade moderna". Não são entidades autônomas, mas construções sociais compreendidas por indivíduos. Decerto, ao nosso ver, antes da percepção de uma comunidade científica, há a constatação de um estatuto social que congrega os indivíduos em torno de uma mesma causa. Para as autoras "a estrutura e organização do conhecimento, os padrões de cooperação, as formas de linguagem e comunicação, os sistemas de informação, a literatura e sua distribuição, os critérios de relevância," são reflexos dos objetos de trabalho destas comunidades, e, também, dos seus papéis na sociedade (HJØRLAND *apud* NASCIMENTO; MARTELETO (2004, [p.5])). Nesse sentido, as autoras propõem características que permitem descrever uma comunidade discursiva, sendo elas:

- a) metas comuns: uma comunidade discursiva tem um conjunto combinado de objetivos, podendo se apresentar em documentos ou em conhecimento tácito;
- b) mecanismos participativos: uma comunidade discursiva tem formas de intercomunicação entre os membros, seja por encontros, correspondência, newsletter ou simples conversas;
- c) troca de informação: a comunidade discursiva usa mecanismos para prover informação com propósitos definidos como, por exemplo, melhora de performance e aumento da capacidade produtiva;
- d) estilos específicos: uma comunidade discursiva usa e possui um ou mais estilos de comunicação para atingir seus objetivos, identificados por tópicos de discussão, forma, posição de elementos e mensagens;
- e) terminologia especializada: uma comunidade discursiva tem um vocabulário

---

extremo oposto, isto é, uma só pessoa detém toda a riqueza. Na prática, o Índice de Gini costuma comparar os 20% mais pobres com os 20% mais ricos.

específico; f) alto nível de especialização: uma comunidade discursiva tem um mínimo de membros com um nível adequado de conhecimento relevante e expertise discursiva. (SWALES *apud* NASCIMENTO; MARTELETO, 2004, [p. 5-6])

Nesse sentido encontramos espaço para aplicar essas descrições de 'comunidades discursivas' à própria prática das humanidades digitais e sua concepção, haja vista a proposta lançada no primeiro capítulo na qual sugerimos a hipótese de "construção de uma narrativa", um "gênero de discurso" identificado em Dalbello (2011). Assim podemos apontar que as nuances dessa estrutura se identifica com o corpo já constituído das *digital humanities* uma vez que 'metas comuns' são recursos altamente representativos para as HDs, pois a essência da área sugere esse movimento no qual indivíduos regidos até então por metodologias tradicionais, passam a almejar que as mesmas metodologias pautem-se ou mediem-se por ferramentas computacionais. Da mesma forma, enquanto 'mecanismos participativos', vimos emergir diversas iniciativas que servem ao propósito da interação, sejam eventos acadêmicos ou espaços qualificados de debates em medida semelhante ao comprometimento com a 'troca de informação', visando fortalecer a interação entre os pares, principalmente por tratar-se de pares oriundos de diversas áreas do conhecimento. Vemos, nesse sentido, enquanto sinalização de tal propósito, a criação de periódicos devotados ao tema e seus pares, à serviço não apenas da divulgação científica, mas fortemente para a convergência e identidade do campo.

O corpo formador das humanidades digitais se desenha mesmo antes de se cunhar a expressão *digital humanities*, estabelecendo, ao nosso ver, seus 'estilos específicos'. Como marco desse processo apontamos o manifesto que "funda" a área. Todavia, esse manifesto, mesmo datando de 2010, é precedido pela publicação de uma obra<sup>12</sup> que já se intitulava com a expressão que, anos depois, se cunharia como termo denominador comum do campo. Ao passo que era introduzida uma expressão responsável por capitalizar as reflexões propostas por uma comunidade discursiva, vemos neste indício, além do esforço de responder questões de uma disciplina emergente, a introdução de uma

---

<sup>12</sup> Nos referimos à obra 'Companion to Digital Humanities (2004)' que deu corpo, ainda em 2004, antes portanto do Manifesto das Humanidades Digitais (2010), às reflexões sobre questões acerca do que seria o campo, suas definições, o que elas incluem ou seus limites porosos, seus integrantes e formadores de opinião, suas metodologias etc SCHREIBMAN (2012). Segundo a autora a obra tornou-se resposta inicial a estas perguntas já que naquela época não havia uma obra referencial que começasse a abordar tais questões, nem havia uma coleção de ensaios que refletissem a interdisciplinaridade e aspectos históricos dessa disciplina emergente. Pelo contrário, havia leitores e práticas em evolução nas interseções da computação com uma única tradicional prática disciplinar. (SCHREIBMAN, 2012, p. [1], tradução nossa)

'terminologia especializada' que passaria a versar entre aqueles que se auto-identificassem com as temáticas da então nova disciplina emergente. Na nossa avaliação a obra citada cumpriu esse papel amplo e complexo de construir 'estilos específicos' em 'terminologias especializadas'.

A última definição da estrutura da comunidade discursiva se aplica ainda mais acuradamente com o contexto das HDs. Decerto temos em sua percepção do 'alto nível de especialização' uma espécie de amálgama que consideramos núcleo convergente dos indivíduos para a formação do campo. Essa motivação, ao nosso ver, centra a comunidade discursiva das *digital humanities* no que as autoras consideram "um mínimo de membros com um nível adequado de conhecimento relevante e *expertise* discursiva". Como este contexto trata-se, basicamente, da aplicação da tecnologia enquanto método, torna-se compreensível a aglutinação de indivíduos não detentores de uma *expertise* em tecnologia em torno daqueles que a detém. Nesse sentido a constatação de um corpo formador do que se identifica como humanidades digitais terá sua expressão mais genuína impressa na produção da literatura proveniente desse campo, versando sobre o tema e suas problemáticas exporão, em nossa avaliação, características singulares deste domínio.

Nesse contexto julgamos o *status quo* das *digital humanities* como quesito apropriado para refletir sobre seu domínio, pois sua incipiência acadêmica nos permite, sem grandes esforços, perceber sua comunidade discursiva de papel determinante na existência do campo. Dessa forma a análise de domínio é produto consequente e lógico dessa identificação via análise de sua produção de literatura, pois segundo NASCIMENTO; MARTELETO, 2004, [p. 6]) reconhecem que as comunidades discursivas se compõem de atores com pontos de vista distintos, estruturas de conhecimento individuais, predisposições, critérios de relevância subjetivos, estilos cognitivos particulares. Para DIAS (2015, p.8) "o domínio é bastante afetado tanto pelas características da comunidade discursiva quanto pela garantia da literatura e demais garantias usadas para validar os instrumentos de representação". Nesse sentido vemos como sequencia causal que o estudo proposto pela pesquisa, ao buscar identificar nas humanidades digitais uma congruência de escopo com a Ciência da Informação usando como estratégia a análise da literatura via uma abordagem bibliométrica, obtém, inevitavelmente, enquanto produto, a persecução de uma comunidade discursiva que nos presta à reflexões no âmbito da análise de domínio.

Todavia, nesse contexto, não podemos deixar de considerar a visão de Smiraglia (2013), segundo ele existem diversas discussões em torno da definição de entidades

variadamente denominados como "domínios, comunidades discursivas e colégios invisíveis que constituem grupos de estudiosos os quais trabalham em problemas de pesquisa que estão, de alguma forma, sendo percebidos como semelhantes" (p.2, tradução nossa). Ou seja, são conceitos utilizados com a finalidade de sugerir algum tipo de rede social entre pesquisadores, todavia com divergências funcionais. Para ele "'domínio' sugere limites intelectuais, 'comunidade discursiva' sugere uma troca ativa de informações; 'colégio invisível' tem sido usado para sugerir tanto a comunidade intelectual como o discurso ativo que ocorrem em um ambiente socialmente estruturado" (SMIRAGLIA, 2013, p. 2, tradução nossa). Para o autor a distinção entre os três conceitos é fundamental, para nossa abordagem nos importa apenas explorar a riqueza contida no insumo fundamental à análise de domínio, que se baseia nas comunidades convergentes de determinado campo enquanto sua matéria-prima inicial. Por outro lado, não obstante, nos preza ressaltar a capacidade de concepção do domínio favorecida por esta abordagem, haja vista sua íntima relação com a ideia de comunidade discursiva. Tanto Smiraglia (2012) quanto Hjørland; Albrechtsen (1995) consideram que o paradigma de análise de domínio dentro de CI tem como foco a descrição de domínios nos quais indivíduos são vistos como membros em participação ativa sob várias maneiras. Assim, para eles, essa participação significa que esses indivíduos fazem parte de uma comunidade discursiva. Dessa forma consideramos, portanto, que uma abordagem sugere reciprocamente a outra. Todavia, apesar de não se pretender determinar as fronteiras do domínio<sup>13</sup> integrado pela comunidade discursiva foco dessa pesquisa, pretende-se enriquecer essa análise produzindo um esboço dessas fronteiras a partir da igual persecução do conteúdo temático produzido por essa comunidade e ventilados através da literatura. Dessa maneira

O trabalho de Hjørland e Albrechtsen não se propõe definir esses termos; em vez disso, sua intenção é demonstrar a utilidade da análise de domínio para o estudo de hipóteses na ciência. Assim, somos levados a **inferir uma relação hierárquica entre um domínio e uma comunidade discursiva**. Hjørland (2002) elabora sobre a metodologia de análise de domínio numa abordagem da ciência da informação, enumerando onze passos que podem fornecer informações sobre um domínio. **Essas etapas incluem estudos bibliográficos e bibliométricos, classificação e indexação**, e várias análises empíricas focadas na terminologia e estruturas intelectuais. (SMIRAGLIA, 2013, p. 2, tradução nossa, grifo nosso).

---

<sup>13</sup> Tennis (2003) *apud* SMIRAGLIA (2013) descreve os limites da definição de "Domínio" e fornece dois eixos para a operacionalização de um domínio. "Esses eixos são 1) **áreas de definições de modulação usadas pelos membros do domínio, que modulam o domínio, afirmando sua extensão**; e, 2) grau de especialização, incluindo "foco" e "Interseção" - que oferecem maneiras de qualificar um domínio aumentando sua intensidade". (SMIRAGLIA, 2013, p. 2, tradução nossa, grifo nosso)"



À medida que consideramos a relação consecutiva entre domínio e sua comunidade de discurso nos baseamos em Smiraglia (2013) e Hjørland (2001) que não negam o caráter empírico concernente à definição - ainda que provisória<sup>14</sup>- de domínio, mas, todavia, ressaltam sua condição impositiva de refletir uma construção social. Assim, para SMIRAGLIA (2013, p. 2, tradução nossa), "a definição de qualquer domínio depende das considerações pragmáticas de seus membros". Discursos, pressupostos teóricos e acordo intersubjetivo são "superiores", ou de correlação mais próxima, quando eles ocorrem dentro de um domínio. Nesse sentido afirmamos, enquanto fundamento de um domínio, que as *digital humanities*, considerando sua comunidade de discurso e seus reflexos teórico-conceituais, reúnem as características concernentes a um domínio, ainda em seu estado atual de incipiente consolidação epistemológica. Assim, segundo SMIRAGLIA (2013, p.2, tradução nossa; grifo nosso), Hjørland profere "uma definição de domínio por **acordo pragmático de seus membros**". Na medida em que eles concordam como componentes, pode-se dizer que eles constituem um domínio. Quanto mais próximo o acordo, maior o grau de concordância de "domínio". Desse modo, consideramos a análise de domínio enquanto ressonância de uma abordagem bibliométrica por sua capacidade de desvelar, a partir da literatura, uma comunidade de discurso.

Essas comunidades de discurso, além de possuírem sua correlação à definição de domínio, também podem ser consideradas enquanto colégios invisíveis e sua formulação conceitual coincide com o *status quo* das humanidades digitais. Para Zucalla (2006) um colégio invisível

é um conjunto de acadêmicos ou cientistas que compartilham interesses de pesquisa sobre uma especialidade, que frequentemente produzem publicações relevantes para determinado assunto e que comunicam formalmente e informalmente uns aos outros para trabalhar em direção a objetivos importantes no assunto, embora possam pertencer à afiliações de pesquisa geograficamente distantes. (ZUCCALA, 2006, p. 155, tradução nossa).

A definição proposta pelo autor para os colégios invisíveis nos permite estabelecer um completo paralelo com a integridade das *digital humanities* celebrada por seu manifesto. Na nossa leitura, o compartilhamento de interesses de pesquisa supõe um distanciamento acadêmico entre os membros do conjunto, pois ao mencionar especialidade ao invés de domínio, há a sugestão desse distanciamento. A iniciativa de

---

<sup>14</sup> Hjørland (2001) não se compromete com uma definição consolidada de 'domínio', dessa forma ele atribui o estado conceitual por ele considerado enquanto uma 'definição provisória'

comunicações formais ou não, na direção a objetivos importantes no assunto, em muito sugere a ideia de comunidades de práticas presente no Manifesto das HDs<sup>15</sup>. Mais uma vez, a utilização do termo 'assunto', supõe que essa comunidade não integra um mesmo contexto ou, pelo menos preliminarmente, um mesmo domínio. Ainda assim, o pertencimento das afiliações de pesquisa geograficamente distantes também nos indigita que o colégio invisível é uma comunidade difusa que encontra nas suas práticas comuns o vínculo que o justifica.

Para SMIRAGLIA (2013, p.3, tradução nossa) os contornos das duas definições de agrupamento sociais são próximas, pois pode-se identificar que "como um domínio, um colégio invisível tem como foco a interação social de um grupo de estudiosos que compartilham interesses de pesquisa semelhantes dentro de uma especialidade de assunto". Para nós essa pontuação se torna crucial no tocante à qualificação daquilo que persecutimos nas HDs, outrora entendido como estrutura elitista da ciência, pois figura-se como maior impositivo do campo o caráter forçosamente convergente dos agentes da temática rumo ao núcleo comum -núcleo esse, inclusive, formador e identitário da disciplina. Todavia Smiraglia (2013) encontra subsídio contundente em Zuccala (2006) para estabelecer que a distinção entre os dois grupos se apresenta, mesmo considerando que ambos também se constituem de "atores sociais engajados no discurso formal e informal" na distinção crucial de que um Colégio Invisível **é relativamente jovem - isto é, sua atividade é na maioria das vezes bastante recente** e isso é visível através de um site dedicado onde os membros podem compartilhar informação de pesquisa atual" (ZUCCALA, 2006 *apud* SMIRAGLIA, 2013, p.3, tradução nossa; grifo nosso).

Assim, nos baseando na contundência da distinção de Smiraglia (2013) e Zuccala (2006), podemos garantir que o corpo de indivíduos atuante nas *digital humanities*, ao passo que reúnem as condições de uma comunidade de discurso reconhecidamente estabelecida e proeminente, também se caracteriza enquanto um Colégio Invisível, por estarem enquadrados na condição de "atividade bastante recente" e tendo nesta constatação sua principal preponderância. Apesar dessa afirmação, os colégios invisíveis podem ser caracterizados por diversos outros fatores, sobretudo o de caráter tecnológico eletrônico que viabilizaria o contato entre membros de uma comunidade reunida à

---

<sup>15</sup> Como versa o Manifesto: 1)"comunidades específicas, oriundas de interesses por diversas práticas, instrumentos ou objetos transversais [...] que convergem atualmente para formar o campo das *digital humanities*"; 2)"[...]atores do *digital humanities*, constituímo-nos numa comunidade de prática solidária, aberta, acolhedora e de livre acesso" (MANIFESTO, 2011, não paginado)

distância<sup>16</sup>. Como no cenário contemporâneo a imposição dos meios eletrônicos de comunicação e socialização tem expandido no contexto social da ciência, consideramos que esta passa a ser uma característica menos representativa de um colégio invisível. Todavia Price (1986) nos concede características outras que permitem pontuar com certa clareza um colégio invisível. Para ele

[...]designa os coletivos informais de cientistas que interagem intimamente, geralmente limitados a um tamanho "que pode ser "invisível", são significativas formações sociais e cognitivas que avançam as frentes de busca da ciência, uma concepção largamente confirmada pelos primeiros estudos de Crane e Mullins [...] Metáforas como o colégio invisível servem para fixar na memória algumas das muitas contribuições para o que se descreve como "o cálculo da ciência". (PRICE, 1986, p. 8-9, tradução nossa).

Podemos identificar na constatação de Price (1986) que a estrutura elitista das humanidades digitais corresponde ao Colégio Invisível da mesma por residir nela o principal núcleo de coletivos informais, pois ao considerarmos a interação entre indivíduos oriundos de diferentes áreas do conhecimento se daria sob uma ótica informal, com os agentes interagindo em meios que não são comuns ou próprios a todos eles. A participação em eventos que não os habituais da área ou a publicação em periódicos também de áreas distintas poderiam configurar formações sociais e cognitivas, não formalmente constituídas e caracterizadas, que avançam as frentes de busca da ciência. Além disso, Price (1986) considera que muitos estudos da sociologia da ciência moderna e dos padrões de comunicação na ciência convergiram para pontuar a característica principal da forma de organização destes grupos e que se tornou conhecido como Colégio Invisível.

---

<sup>16</sup> Autores como MOREIRA (2005) sugerem o termo "Colégio Virtual" para caracterizar a principal ação de relacionamento praticada entre os membros de um colégio invisível. Segundo ele "o conceito em formação de colégios virtuais é apresentado muitas vezes como sucedâneos dos colégios invisíveis. A terminologia neste domínio reflete a velocidade das mudanças tecnológicas a que está afeito. Enquanto o termo colégio invisível representou por quatro décadas um conceito relativamente consensual, as novas designações são legião, se não conflitantes. Diferenças sutis podem ser verificadas na lista de denominações apresentada por Gresham Jr. (1994, p. 39): *electronic conferences, e-conferences, computer conferences, mailing lists, listservs, electronic forums, online discussion groups, scholarly discussion groups, special interest groups, news groups, netgroups*. Esta taxionomia tem, como se vê, dez anos, e os conceitos ainda não se firmaram totalmente. Adotar-se-á aqui o termo colégio virtual com o fim de preservar a analogia com o colégio invisível e sua carga semântica, conforme proposto por Price (1973), isto é, tem-se como princípio que o número de cientistas trabalhando em uma área específica é pequeno e que estes se conhecem entre si, mesmo que não pessoalmente. Esta comunidade mantém-se a par dos respectivos trabalhos e troca informações via correio eletrônico ou listas de discussão, preferencialmente, embora utilize também outros meios. O colégio virtual atua como rede de comunicação e intercâmbio, como fórum de educação e de socialização dos novos cientistas". (MOREIRA, 2005, p. 58)

Tanto a forma de organização quanto seus padrões de comunicação balizariam, para além do fato da interação entre autores e pelo seu envolvimento na colaboração mútua, o fornecimento de uma técnica susceptível de concordância com as ligações de citação e com estimativas subjetivas que também podem se constituir enquanto sintomas de um colégio invisível. Nesse sentido vemos nas relações sociais constituidoras das características dos colégios invisíveis as interações de ordem bibliográficas, assim como Price (1986) que atribui aos dados de natureza bibliográfica e autorias colaborativas a validade de servir como indicações de links que se mostram úteis como um meio de analisar colégios invisíveis.

Por certo as considerações em torno da existência de um colégio invisível no âmbito das *digital humanities* corrobora a proposta de analisar esse domínio através da relação entre seus agentes. Essa relação, a despeito do formalismo social ou acadêmico, pode ser monitorada por intermédio da análise dos protocolos bibliométricos que, apesar de não se constituírem os únicos, são os que se melhor apresentam enquanto modelos que constituíram nossa escolha para uma análise metodologicamente científica. Por fim, nos cabe enquanto instrumento de identificação e mensuração da hipótese por essa pesquisa defendida, de uma verificação da participação contundente de agentes provenientes ou integrantes da área informacional como partícipes desse Colégio Invisível que para nós se constitui a estrutura elitista da ciência no domínio das humanidades digitais.

### **3.3 A centralidade dos indicadores de mensuração do conhecimento: a formação da estrutura elitista da ciência**

O cenário da sociedade da informação fomentado também pelo produtivismo acadêmico como um de seus sintomas e/ou causas desse fenômeno, delega importância ainda maior aos estudos bibliométricos. Esse movimento, ou mesmo evolução das abordagens em torno desses estudos, se deu pela necessidade de refinamento da mensuração da produção acadêmica. Com a imponência do mundo digital e a consecutiva migração dos meios de veiculação da informação do físico para digital/virtual colocou aos estudos métricos a imposição de servir de base para mensuração da informação em suas novas formas de espalhamento. Nesse contexto vimos surgir abordagens dinamizadas para abrangerem uma mensuração da informação em seu "estado" menos

específico<sup>17</sup>, à ciência<sup>18</sup> e mais recentemente a web<sup>19</sup>. Para VANTI (2002, p. 152) "estas subdisciplinas, apesar de apresentarem algumas semelhanças ou pontos de convergência, possuem características, enfoques e funções dissímeis". Todavia a amálgama que une essas manifestações de mensuração do conhecimento é a mesma que fundamenta sua importância, pois num cenário de produção exponencial de conhecimento em escala inédita exige-se uma maneira de averiguar a qualidade do produto informacional. Assim, para a autora

A avaliação, dentro de um determinado ramo do conhecimento, permite dignificar o saber quando métodos confiáveis e sistemáticos são utilizados para mostrar à sociedade como tal saber vem-se desenvolvendo e de que forma tem contribuído para resolver os problemas que se apresentam dentro de sua área de abrangência. (VANTI, 2002, p.152)

Assim, temos na avaliação de, outrora a literatura e, agora, nas mais diversas formas de vetorização do conhecimento, a construção da relevância de agentes, atores e da literatura mediante as análises sobre esse fenômeno nesses contextos. Como consequência o movimento passa a jogar luz também sobre a relação entre tais indivíduos,

---

<sup>17</sup> Fazemos, aqui, alusão à Informetria, que segundo VANTI (2002, p.154-155) "seu surgimento é posterior ao dos outros dois termos, pode-se dizer que esta tem um escopo tanto mais distinto e abrangente do que a Cienciometria e a Bibliometria, pois, ainda de acordo com a conceituação de Tague-Sutckiffe, glossada por Macias-Chapula. [...] Informetria é o estudo dos aspectos quantitativos da informação em qualquer formato, e não apenas registros catalográficos ou bibliografias, referente a qualquer grupo social, e não apenas aos cientistas. A informetria pode incorporar, utilizar e ampliar os muitos estudos de avaliação da informação que estão fora dos limites da bibliometria e cienciometria. O termo informetria designa, conforme Hjotgaard Christensen & Ingwersen, uma extensão recente das análises bibliométricas tradicionais ao abarcar o estudo das modalidades de produção da informação e de comunicação em comunidades não acadêmicas". Para estes autores, as análises de difusão de determinados assuntos nos bancos de dados full-text dos jornais seriam uma das novas possibilidades que surgem neste campo. A informetria se distinguiria claramente da cienciometria e da bibliometria no que diz respeito ao universo de objetos e sujeitos que estuda, não se limitando apenas à informação registrada, dado que pode analisar também os processos de comunicação informal, inclusive falada, e dedicar-se a pesquisar os usos e necessidades de informação dos grupos sociais desfavorecidos, e não só das elites intelectuais.

<sup>18</sup> Nos referimos à Cienciometria, termo que segundo VANTI (2002, p.153), "surgiu na antiga URSS e Europa Oriental e foi empregado especialmente na Hungria. Originalmente, referia-se à aplicação de métodos quantitativos para o estudo da história da ciência e do progresso tecnológico. As primeiras definições consideravam a cienciometria como "a medição do processo informático", onde o termo "informático" significava "a disciplina do conhecimento que estuda a estrutura e as propriedades da informação científica e as leis do processo de comunicação". Este termo alcançou notoriedade com o início da publicação, em 1977, da revista *Scientometrics*, editada originalmente na Hungria e atualmente na Holanda".

<sup>19</sup> Para VANTI (2002, p.156), "as análises quantitativas, com os avanços tecnológicos, veem-se facilitadas e ao mesmo tempo encontram novos e estimulantes campos de atuação. É o caso dos estudos que estão sendo desenvolvidos atualmente sobre o conteúdo e a estrutura das home-pages na Web. De acordo com Almind & Ingwersen, está nascendo uma nova área de interesse dentro da informetria: a webometrics. A webometrics ou webometria consiste, conforme a definição esboçada por estes últimos autores, na aplicação de métodos informétricos à World Wide Web. Nesta mesma linha, Cronin & McKim têm destacado que a Web está se tornando, cada vez mais, um importante meio de comunicação para a ciência e a academia, pelo qual é lógico que os estudos quantitativos se estendam também a este ambiente".

pois reside na avaliação e sua consecutiva aceitação pelos pares daquilo que se configura como principal instrumento de validação da qualidade e a consecutiva proeminência de determinadas produções científicas. É nesse sentido que esse esforço volta-se a observar e atribuir substancialidade ao movimento de relacionamento social na ciência caracterizado principalmente pela citação entre pares. Dessa forma podemos aplicar a conceituação de rede social que Newman (2001, p.1, tradução nossa), define como "um conjunto de pessoas, cada uma das quais é familiarizado com algum subconjunto dos outros. Tal rede pode ser representada como um conjunto de pontos (ou vértices/nós) denotando pessoas, unidas em pares por linhas (ou arestas) denotando conhecimento". Assim deve ser factível visualizar uma rede de atores proeminentes a serem caracterizados enquanto uma estrutura elitista da ciência que, por sua vez, fora construída através das relações entre os indivíduos pela validação como tal, fruto da relevância de suas publicações de uns para os outros. Nesse sentido

Uma rede de coautoria entre pesquisadores, também chamada de rede de colaboração científica, é um exemplo particular de uma rede social. Neste caso, os atores são autores de artigos e um laço existe entre dois autores se eles já escreveram algum artigo juntos. **A produção de um artigo pode ser considerada como uma forma de documentação da colaboração entre dois ou mais autores [NEWMAN 2004]. Assim, a rede de coautoria de uma comunidade pode revelar fatos interessantes a seu respeito, como, por exemplo, a existência de grupos que colaboram mais densamente, a identificação de relacionamentos mais intensos entre determinados autores ou ainda autores com maior grau de colaboração.** O estudo de redes de coautoria pode ser utilizado também para se comparar os padrões de colaboração entre diferentes comunidades científicas. (PROCÓPIO JÚNIOR; LAENDER; MORO, 2011, p. 1, grifo nosso)

Para Newman (2004) essas colaborações em forma de coautoria formam redes que permitem visualizar particularidades desses complexos sistemas. Para o autor a percepção das redes de coautoria de artigos em periódicos fornece, por exemplo, uma janela sobre padrões de colaboração dentro da comunidade acadêmica. Assim, segundo ele, "a rede de coautoria é tanto uma rede representando a sociedade acadêmica como é uma rede retratando a estrutura do conhecimento" (Newman, 2004, p.1). Nesse sentido esta pesquisa objetivou atingir essas redes de estrutura da ciência através da literatura indexada por humanidades digitais. A produção científica extraída das bases de dados mais abrangentes e consolidadas como Scopus e Web Of Science, nos serviu de matéria-prima para a construção dessas redes sociais do domínio das *digital humanities*. A partir das análises das métricas da bibliometria como, por exemplo, de citação, cocitação, coautoria, acoplamento bibliográfico etc, e sua aplicação ao nível de autores, periódicos

e palavras-chaves nos revelam um cenário satisfatório sobre o domínio das humanidades digitais.

Assim sendo, o próximo capítulo se prestará a apresentar os dados extraídos das fontes de literatura, a construção dessas redes, a metodologia aplicada partindo da construção dos termos para revisão de literatura, das métricas aplicadas ao VOSviewer para refinamento e composição da estrutura elitista que buscamos evidenciar para, nela basear nossas análises-, além das análises sobre esses resultados numa clara tentativa de esboçar o domínio das *digital humanities* além de conhecer seu arcabouço teórico-conceitual proveniente da análise dos termos indexadores.

#### **4 DESVELANDO AS HUMANIDADES DIGITAIS A PARTIR DA LITERATURA: UM ENSAIO PARA COMPREENSÃO DO DOMÍNIO EM RESSONÂNCIA À ABORDAGEM BIBLIOMÉTRICA**

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 em análise. Assim, nesse contexto, ganha corpo a análise de citações: um expoente dos estudos bibliométricos que atualmente serve de base para construção de diversas redes à serviço da compreensão do conhecimento de um determinado campo.

A análise de citações é uma técnica bem conhecida que tem sido aplicada em vários campos de pesquisa para estudar, entre outros, os fluxos de conhecimento, difusão de idéias, estruturas intelectuais da ciência, relevância dos recursos de informação e avaliação de pesquisadores e instituições de pesquisa. Entre os campos de pesquisa que empregaram métodos de análise de citações, a sociologia, a história da ciência, a biblioteconomia e a ciência da informação, a ciência gerencial e a política de pesquisa são as mais proeminentes. Juntamente com a indexação de citações e a vinculação de citações, a análise de citações também fornece as bases para uma recuperação de informações eficaz que, aplicada a links da Web, estava no centro do sucesso do mecanismo de pesquisa do Google. (ZHAO; STROTMANN, 2015, p. 1, tradução nossa) .

Não restrita ao campo dos estudos informacionais, a relevância desses estudos fomentou sua expansão encontrando pertinência em diversas outras áreas do conhecimento e por campos científicos como a Sociologia e a História da Ciência, por exemplo. Para os autores houve um crescente interesse na análise de citações como auxílio à abordagem de várias questões de pesquisa, gerenciamento ou serviço de informações, como classificações de universidades, avaliação de pesquisas e visualização de domínio do conhecimento. Dessa forma os autores identificam que os "dados de citação digital e o poder do computador tornaram possíveis os estudos de análise de citações em grande escala e resultaram em novos desenvolvimentos interessantes em fontes de dados, bem como técnicas e ferramentas para coleta de dados de citações, análise e visualização" (ZHAO; STROTMANN, 2015, p.1, tradução nossa).

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 são representados 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, e um cientista como aquele que escreve artigos científicos. Logo a expressão do conhecimento registrado e produzido através dos artigos científicos alcança, enquanto exponência da evolução dos saberes, sua legitimação.

A relevância dessas análises baseadas em citação concentram-se na percepção de que cada trabalho ou pesquisa que cita uma fonte anterior acaba por ter nela, indissociavelmente, uma parte e/ou característica daquela fonte inspiradora. O que autores no passado consideravam como a disseminação de uma doença, Zhao; Strotmann (2015) sintetizam

A lista de referência em um trabalho de pesquisa é uma parte essencial do artigo. Ao apontar publicações anteriores que influenciaram a pesquisa relatada no presente trabalho de uma forma ou de outra, as referências ligam o artigo atual a essas publicações anteriores e, por extensão, à rede global de publicações de pesquisa. É geralmente assumido que **uma citação representa o uso de citar autores do trabalho citado, e indica uma influência do trabalho citado sobre os autores do novo trabalho e, como tal, um fluxo de conhecimento do citado para os autores das obras citantes.** As citações também indicam parentesco entre esses dois trabalhos. (ZHAO; STROTMANN, 2015, p. 1, tradução nossa, grifo nosso).

Desta sorte, por sua relevância, a análise de citações serve à ciência como instrumento de investigação de seu próprio desenvolvimento correspondendo a uma parte indissociável do processo de construção do conhecimento. Nesse sentido, assertivamente, esses estudos fundamentaram a heurística que tange os estudos sobre o domínio tendo também no âmbito destes estudos um instrumento capaz de servir aos contornos intrínsecos do desenvolvimento dos campos científicos, suas fronteiras e características de pertencimento. Para os autores é importante entender o fundamento dos estudos de análise de citação clarificando o processo de citação como estando caracterizado enquanto o movimento de "quando o artigo 'A' faz referência ao artigo 'B', costuma-se dizer que 'A' cita ou referencia 'B', e 'B' é citado por, recebe uma citação de, ou é uma das “referências citadas” em 'A'. Em essência, uma referência do artigo 'A' ao artigo 'B' é uma citação recebida por 'B' de 'A'." (ZHAO; STROTMANN, 2015, p. 2, tradução nossa).

Para De Bellis (2009) a análise de cocitação estaria a serviço de uma cartografia científica, configurando segundo ele uma estrutura especial da ciência. O autor justifica que a análise de documentos altamente cocitados viabilizam o esclarecimento de algumas

questões como "Quais são as unidades estruturais naturais da ciência?" ou "Como essas unidades estruturais se relacionam umas com as outras? Elas estão frouxamente ou firmemente integradas, ou completamente isolados um do outro?" ou também "Quais são as forças que determinam essas unidades estruturais e suas inter-relações? Até que ponto o mapa da ciência é também um mapa do conhecimento científico?" e ainda "Qual o papel dos fatores sociais na determinação da estrutura? Como a estrutura da ciência muda ao longo do tempo, tanto a nível macro como a nível micro?" (DE BELLIS, 2009, p. 156, tradução nossa). A nosso ver essas análises fornecem subsídios amplos embasando, potencialmente, mesmo análises mais profundas e de ordem epistemológica. Apesar de não nos propormos a uma análise dessa profundidade no domínio em questão, corroboramos dessa forma, a legitimidade destes estudos para uma análise suficientemente reveladora sobre um determinado domínio.

Nisto posto, para os autores, em especial Zhao; Strotman (2009), suas funções à serviço da ciência podem ser pontuadas

Em geral, a análise de citações pode auxiliar efetivamente na descoberta de novos conhecimentos e no gerenciamento e uso de recursos de conhecimento existentes (Garfield, 1979; Swanson, 1986; Small, White et al., 2000). Em particular, a análise de citações pode ser usada para; 1. avaliar recursos de informação e avaliar contribuições acadêmicas, 2. **mapear campos de pesquisa para estudar suas estruturas intelectuais**, 3- rastrear fluxos de conhecimento e difusão de idéias; estudando usuários e usos da literatura acadêmica, e organização, representação e recuperação. (ZHAO; STROTMANN, 2015, p. 2, tradução nossa, grifo nosso).

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 De Bellis (2009) essa indicação cognitiva já fora pensada, no passado, em função de servir a algoritmos de indexação, devido ao seu potencial cognitivo, salientando que essa técnica poderia melhorar o desempenho de um sistema de recuperação de informação, não sendo surpreendente que "bancos de dados e bibliotecas digitais incorporassem sistematicamente alguma forma de acoplamento bibliográfico nos algoritmos projetados especificamente para recuperar documentos ou páginas da Web" (DE BELLIS, 2009, p. 157, tradução nossa). Esse caráter cognitivo destacado pelo autor nos embasa também no sentido de cotejar a carga cognitiva fruto dessas análises, com sua função equivalente na Ciência da Informação.

Dessa forma, baseados nessa premissa de que a vinculação de trabalhos e autores confere uma estrutura reveladora de um domínio, pretendemos, a partir de agora, identificar que universo compreende nossa análise, bem como que recorte esta pesquisa aplicou além de, principalmente, apresentar as metodologias que, previamente, nos permitiram constituir matéria-prima para compor o objeto de análise e, posteriormente aquela que aplicamos para modelar o universo analisado. Dessa forma este capítulo se presta a apresentar o *corpus* analisado, sua metodologia de constituição enquanto nossa matriz de dados, identificar as métricas de coautoria (autores, organizações e países), citação (referências, periódicos, autores, organizações e países), cocitação (referências citadas, periódicos citados e autores citados), coocorrência de palavras-chave e acoplamento bibliográfico (referências, periódicos, autores, organizações e países).

Acrescentaremos à análise das métricas o recurso de visualização de tais cenários, dessa forma a apresentação destes pôde ser concebida visualmente, agregada de sua rede de relacionamentos entre termos. Assim a proposta deste capítulo é, a despeito dos métodos tradicionais de análises dessa natureza, não apenas apresentar os identificadores bibliométricos, mas dar corpo à sua existência via construção e apresentação de grafos. Dessa forma delegamos ao *software* VOSviewer a função de construção desses recursos visuais e, principalmente, autonomia integral para cálculo da relevância dessas métricas. Assim, nosso balizamento conceitual está todo fundamentado na teorização bibliométrica, via aplicação de conceitos próprios a esse contexto, porém relegando a importância dos cálculos de natureza matemática à segundo plano, dando relevância à visualidade gráfica dos mesmos.

#### **4.1 Constituindo uma matriz de dados para análise: recuperando a literatura sobre humanidades digitais**

A análise proposta por esta pesquisa está baseada na investigação da literatura produzida e indexada por 'humanidades digitais' nas bases de dados consolidadas globalmente e consideradas instrumentos agregadores de artigos científicos de natureza multidisciplinar. Nesse sentido a pesquisa se preocupou em abarcar bases de dados que sirvam à ciência enquanto repositórios informacionais dedicados a reunir o conhecimento produzido e disseminado nas mais diversas áreas de conhecimento. Mais que isso, nossa intenção foi utilizar repositórios que estivessem legitimados como fontes com amplitude

ímpar para agregação de artigos científicos, ou seja aquelas que desfrutem de legitimação como as mais abrangentes em termos de volume recuperável.

Nesse sentido nossa escolha recaiu sobre duas das mais consolidadas<sup>20</sup> no meio acadêmico brasileiro: Web Of Science<sup>21</sup> e Scopus. As bases foram selecionadas por estarem caracterizadas como bases multidisciplinares, essa peculiaridade é fundamental à pesquisa por minorar a incidência de vieses caso o levantamento se utilizasse de bases destinadas a algum nicho de conhecimento. Ao nosso ver essa questão imputaria um viés nos resultados por refletirem a produção contextualizada na área respectiva, por essa razão bases específicas não se habilitam ao levantamento proposto.

Uma das bases mais disseminadas, a Web Of Science, foi desenvolvida pela Thomson Scientific, parte da Thomson Corporation e “tem dominado o campo de referência acadêmica, principalmente através da divulgação anual de o fator de impacto dos artigos, uma ferramenta para avaliar importância e influência de publicações específicas” (FALAGAS et al, 2008, p. 341, tradução nossa). O autor destaca que sua relevância cresce à medida que o sistema de medição de impacto das publicações se populariza, o que confere importância às publicações mais acessadas, além disso teve sua

---

<sup>20</sup> Por consolidadas nos referimos àquelas que maior presença e popularidade possuem no contexto acadêmico. Esse contexto deve ser entendido dentro da percepção de que esses repositórios não são de livre acesso, dependendo, portanto, de uma estrutura que possibilite tanto a oferta quanto o acesso. Nesse caso pontuamos que as bases de dados disponíveis no Brasil estão abarcadas dentro do que se conhece por Portal de Periódicos Capes. Segundo o próprio Portal de Periódicos ele tem o objetivo de fortalecer a pós-graduação no Brasil, remontando o ano de 1990 quando o Ministério da Educação criou em o programa para bibliotecas de Instituições de Ensino Superior. A partir dessa iniciativa, cinco anos mais tarde, foi criado o Programa de Apoio à Aquisição de Periódicos. O Programa está na origem do atual serviço de periódicos eletrônicos oferecido pela Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (Capes) à comunidade acadêmica brasileira. O Portal de Periódicos foi oficialmente lançado, da forma que é conhecido atualmente, em novembro de 2000, na mesma época em que começavam a ser criadas as bibliotecas virtuais e quando as editoras iniciavam o processo de digitalização dos seus acervos. Com o Portal, a Capes passou a centralizar e otimizar a aquisição desse tipo de conteúdo, por meio da negociação direta com editores internacionais. Com a centralização desse processo de aquisição e disponibilização desse conteúdo pela CAPES a grande maioria do acesso da comunidade acadêmica brasileira se dá relativamente padronizada por esse mecanismo. Nesse sentido nos convém pontuar que, do portfólio oferecido à pesquisa acadêmica nacional, Web Of Science e Scopus são as mais abrangentes nesse contexto e no segmento "foco multidisciplinar".

<sup>21</sup> O Web of Science (anteriormente conhecido como Web of Knowledge) é um serviço de indexação de citações científicas on-line baseado em assinatura originalmente produzido pelo Institute for Scientific Information (ISI), agora mantido pela Clarivate Analytics (anteriormente propriedade da Intellectual Property and Science da Thomson Reuters), que fornece uma pesquisa abrangente de citações. Ele dá acesso a vários bancos de dados que fazem referência à pesquisa interdisciplinar, que permite a exploração em profundidade de subáreas especializadas dentro de uma disciplina acadêmica ou científica. (WEB of Science, Wikipedia: A Enciclopédia livre. Disponível em <[https://en.wikipedia.org/w/index.php?title=Web\\_of\\_Science&oldid=856456072](https://en.wikipedia.org/w/index.php?title=Web_of_Science&oldid=856456072)> Acesso em: 7 set 2018).

origem a partir do Science Citation Index<sup>22</sup>. Já a Scopus, segundo FALAGAS et al (2008, p. 339, tradução nossa) foi desenvolvida pela Elsevier, “combinando as características da PubMed e da Web of Science”. Estas características combinadas permitem maiores funcionalidades, tanto para a pesquisa em literatura médica quanto as necessidades acadêmicas em geral como, por exemplo, análise de citações. Nesse sentido essas condições conferem às bases de dados o papel de principais centralizadoras do conhecimento acadêmico registrado, veiculado e acessível, viabilizando, portanto as práticas de levantamento bibliográfico base dos estudos bibliométricos. Assim por reunir tais condições além de abrangência, capacidade de indexação e disponibilidade ampla, as bases de dados eletrônicas passaram a compor o processo de revisão de literatura enquanto condição *sine qua non* para legitimação desse método baseado, entre outras qualidades, na possibilidade de superação de limites naturalmente impostos pelos sistemas não eletrônicos e colaborativos praticados anteriormente ao advento da internet.

De toda forma este estudo privilegiou a base Scopus por ser esta aquela que mais itens reporta em seu repositório. Esta escolha se fez necessária, pois existe uma sobreposição severa de artigos registrados nas duas bases. Essa sobreposição da ordem de maior que 50%<sup>23</sup>, imprimiria ao estudo uma redundância desnecessária e prejudicial às análises. Dessa maneira justificamos nossa escolha pelos resultados exclusivamente reportados pela base de dados Scopus.

As revisões de literatura que possuem um papel fundamental para ciência, usufruíram de um aperfeiçoamento considerando a disponibilidade de bases de dados eletrônicas -fenômeno que reconfigurou essas práticas, por permitir não somente a atualização de uma dada problemática, mas também por viabilizar uma percepção de seu panorama histórico. Para FIGUEIREIDO (1990, p. 132) as revisões de literatura "constituem-se em parte integral do desenvolvimento da ciência enquanto função histórica; e fornecem aos profissionais de qualquer área, informação sobre o desenvolvimento corrente da ciência e sua literatura enquanto função de atualização".

---

<sup>22</sup> Para Eugene Garfield criador do Science Citation Index "essa iniciativa foi a primeira tentativa consistente de compilação de bibliografia universal para controle da literatura científica desde a virada do século, uma forma de compilação de um inventário da ciência - precisamente o que o Science Citation Index é". (GARFIELD, 1984, p. 525 tradução nossa).

<sup>23</sup> Analisamos conjuntamente as duas matrizes de dados reunindo-as no software Zotero para identificar ocorrências iguais (duplicadas) de artigos. Como essa verificação alcançou, em determinado momento do processamento, a metade das ocorrências da base com maior número de artigos, escolhemos descartar a base com menor número de ocorrências (Web Of Science) e basear a análise exclusivamente no repositório com maior número de itens (Scopus).

Para a autora as funções histórica e de atualização, como matéria-prima à serviço das análises sobre o resultado dessas revisões, podem ser pontuadas em inúmeras colocações:

Funções históricas: 1. Comparação de informação de fontes diferentes - já que os resultados de muitos trabalhos são agrupados para **oferecer uma visão geral do estado corrente do desenvolvimento de um assunto, em um dado tempo. Fornece assim uma base conceitual, um processo unificador do conhecimento**, do qual trabalhos individuais podem ser vistos em perspectiva. 2. Compactação do conhecimento existente. 3. Substituição dos trabalhos originais como o registro escrito - já que das revisões são extraídos os registros incluídos em livros e no corpo geral do conhecimento estabelecido. 4. Identificação de especializações emergentes. 5. Direcionamento da pesquisa para novas áreas. As funções de atualização destes artigos são: 1. Notificação informativa sobre a literatura publicada, permitindo ao cientista concentrar-se nos trabalhos mais valiosos. 2. Serviço de alerta para campos correlatos, tornando o cientista ciente do que ocorre em outras áreas de menor interesse e que, portanto, não requerem a leitura do original. 3. Suporte à busca bibliográfica, assegurando que todos os documentos importantes foram levados à atenção do leitor. 4. **Orientação inicial para uma nova área, já que a combinação de informação didática e bibliográfica torna as revisões um veículo ideal para fornecer informação básica para o pesquisador iniciando um novo projeto ou mudando para nova direção.** 5. Auxiliar no ensino, de maneira semelhante a anterior, e particularmente para alunos envolvidos em trabalhos acadêmicos. 6. Feedback, fornecendo uma avaliação razoavelmente objetiva do valor do trabalho publicado. (FIGUEIREDO, 1990, p. 132-134, grifo nosso).

A autora garante a abundância dessas funções à serviço da análise sobre o produto obtido de uma revisão de literatura e, contudo, duas dessas funções estão diretamente ligadas ao que esta pesquisa se propõe que foi obter uma visão geral corrente de determinado assunto e o fornecimento de uma base conceitual respectiva a esse assunto. Além disso, a autora nos respalda conforme a aplicação da revisão enquanto método deste estudo ao identificar como funções de atualização "a orientação inicial para uma nova área já que a combinação de informação didática e bibliográfica torna as revisões um veículo ideal para fornecer informação básica para o pesquisador" (FIGUEIREDO, 1990, p. 133). A autora ainda assinala que os artigos responsáveis por publicar tais revisões de literatura podem ser considerados como sendo o mesmo que 'artigo de estado da arte'; estas publicações, por sua vez, são também conhecidas como 'relatórios do estado da arte', ou ainda como *progress reports* ou relatórios de avanços em determinada área (PURCELL; SCHIACHTER *apud* FIGUEIREDO, 1990, p. 134). Com efeito, podemos considerar a pujança da contribuição dos estudos de revisão de literatura para a ciência em geral, não tendo sua legitimação circunscrita a campos peculiares como, por exemplo, o da Ciência da Informação.

Nesse sentido a revisão de literatura nos fornece um panorama teórico do tema. Como defende BARONE (1990, p. 138), a revisão "procura um estatuto teórico que possa

valer como referencialidade diante das várias/mesmas posturas descritivas, metodológicas e técnicas diante do objeto, por parte dos textos que compõem a literatura". Essa assertividade, a nosso ver, agrega à nossa proposta de análise da literatura, pois busca, através das medidas de relevância, localizar o núcleo mais representativo do produto recuperado pela revisão. O autor associa a revisão de literatura a um texto científico que tem por objeto "um conjunto de textos que, por vezes, são construções passíveis de visão analítica individualizada, mas constituem, acima de tudo, partes do conjunto maior que a revisão de literatura" (BARONE, 1990, p. 138). Ou seja, apesar da revisão estar centrada numa temática, é natural que parte de seus achados não estejam em sintonia. Todavia, o método de atribuição quantitativa de relevância praticado pela ferramenta utilizada por esta pesquisa é capaz de contornar essa questão, tendo condições de processar e obter uma espécie de sumo do universo recuperado.

Muito embora o método de revisão de literatura tenha adquirido legitimação na área da Ciência da Informação, o processo consolidante dessa prática se deu baseado na instrumentalização de tal método. Indica BARONE (1990, p.138) que se tende a chamar de teoria "um conjunto de técnicas fundadas em princípios intocáveis tais como aquele da utilidade da informação a priori". Para o autor, a revisão de literatura carece de uma teoria baseada, segundo ele, numa sistematicidade propositiva por estar focada num 'rebotinho puramente pragmático'. Para ele, portanto, não parece tão temerário assim dizer que a revisão de literatura fundada nas estruturas de representação dos textos, "se levada a cabo de forma sistemática, poderá limpar a área desse rebotinho puramente pragmático e redirecioná-la em busca de uma teoria de fato" (BARONE, 1990, p. 138). Assim, Barone (1990) nos parece visionar, o que se estabeleceu -importante que se diga fora da área da informação- como modelo sistemático de revisão de literatura

o desenvolvimento dessa questão é vital para a área pois garante o surgimento de constructos teóricos básicos dos quais a área, até hoje, sofre a falta. Sabemos que a ênfase ao pragmatismo, tornada prática histórica de uma área do conhecimento, transforma esta última em pura matéria instrumental destituída de essência científica. (BARONE, 1990, p.139).

A percepção do autor atenta para a carência de teoria científica como um ponto nevrálgico dos estudos desse gênero e, ao nosso ver, como resposta ao surgimento de modalidades especializadas para as revisões de literatura. Essa especialização das revisões teve na área médica sua maior ressonância, consolidando, portanto, as revisões sistemáticas de literatura, modalidade que se estabeleceu enquanto evidência científica. De acordo com SOUZA; RIBEIRO (2009, p. 241) "a revisão planejada da literatura

científica usa métodos sistemáticos para identificar, selecionar e avaliar criticamente estudos relevantes sobre uma questão claramente formulada”. Com objetivos empiricamente definidos, a busca por respostas na literatura para problemas específicos concedeu ao método *status* de evidência científica. A constância de problemas práticas lidadas na área médica favorece esses estudos por provocarem problemas e soluções sensíveis, fazendo com que os levantamentos apresentem respostas concretas. Pontuam os autores que o objetivo da sistematização é “reduzir possíveis vieses que ocorreriam em uma revisão não-sistemática, tanto os vieses observados na forma de revisão da literatura e na seleção dos artigos quanto aqueles detectados pela avaliação crítica de cada estudo” (SOUZA; RIBEIRO, 2009, p. 241).

Para Cooper; Hedges (2009) a revisão sistemática da literatura trata de uma agregação de configurações peculiares das características de revisões de literatura, baseado em estudos primários passíveis de organização de fatores cronológicos, conceituais e metodológicos. Direcionando-se a públicos específicos, pode ser considerada ou entendida como uma análise das análises, se configurando, portanto como uma meta-análise. Nesse método da revisão sistemática da literatura se confere à estrutura um alto rigor metodológico possibilitando ou objetivando a análise de estudos já realizados, atualizando-os ou dando início a estudos primários.

A revisão sistemática da literatura possui características que a permite diferenciar do método de revisão tradicional, apesar de ambas terem como objetivo um levantamento de literatura produzida. Assim, em linhas gerais, no método sistemático foca-se na informação primária enquanto que, na tradicional, em autores emblemáticos. Apesar disso, ambas se baseiam em varredura de literatura passada na expectativa de definir um *status quo* para a pesquisa pretendida. No caso da revisão tradicional essa expectativa se baseia em literatura referencial ou, como nas palavras de Cooper; Hedges (2009), na busca de reunir literatura passada que se acredita relacionar com um assunto em comum. Outra característica apontada que diferencia as duas revisões é o fato de que no processo tradicional não necessariamente agrega-se conclusões ou a abrangência da cobertura dos trabalhos. Nesse modelo não há uma preocupação primordial com a metodologia de trabalho, dificultando assim, uma revisitação do processo. Para o modelo sistemático, há uma fundamentação metodológica crucial que permite, entre outras questões, a revisitação daquele processo particular por utilizar um balizamento rigoroso dos métodos.

A revisão sistemática da literatura possui, por exemplo, estágios que oferecem uma estrutura conceitual das pesquisas sistemáticas. O levantamento parte da formulação



do problema, que deve amparar-se em uma pesquisa primária prévia na expectativa de mensuração de um volume de pesquisas existentes sobre o problema, que variará de acordo com as características desse problema. No estágio da busca de literatura é importante deter-se a um alvo que orientará a uma cobertura exaustiva das pesquisas conduzidas no tópico de interesse. Os conteúdos coletados nesses estágios iniciais são então submetidos nos estágios de avaliação e análise dos dados. Esses estágios devem se deter à identificação de partes dos documentos que auxiliam na resposta do problema, extrair a literatura relevante, identificar divergências ou seja, o objetivo dessas etapas seriam obter um modelo de variáveis numerosas e suas inter-relações. Por fim os estágios de interpretação dos resultados e apresentação desses resultados, em que no primeiro se deve estimar e interpretar os efeitos dos pontos de moderação entre suas variáveis e, no segundo, estabelecer critérios para tornar público o contexto da pesquisa, seus métodos, resultados, além do sentido da revisão sistemática para a pesquisa em questão e os achados do desafio final da síntese.

Norteados pelas diretrizes praticadas pelas revisões sistemáticas, todavia cientes de que sua aplicação não se viabilizaria neste levantamento por nos imputar um recorte prévio, que deixaria de abarcar parte importante da literatura para análise, conduzimos a elaboração da expressão de busca pela preocupação de fazê-lo minimizando vieses ao mesmo tempo que almejando contemplar o universo mais abrangente de literatura indexada como *digital humanities*.

#### 4.1.1 Elaboração da expressão de busca

Partindo dessa proposta, elaboramos uma expressão de busca para dar conta de recuperar a publicação sobre *digital humanities* em inglês, espanhol e português para proceder à pesquisa nas base de dados eleitas para esta revisão. Dessa forma a expressão de busca desenhada para esse levantamento se registra em torno da carga conceitual dos seguintes termos: **1-digital humanities**, **2-humanities computing**, **3-humanidades digitais**, **4-humanidades digitais**, além de **5-digital research**, **6- digital research activities** e, por fim, **7- digital methodology**. O primeiro termo foi enriquecido por *string* na expressão de busca visando abranger variações como, por exemplo, o termo no singular; o segundo termo dá conta da pré-nomenclatura da área, já que antes de alcançar

sua denominação oficial tem cunhado nessa expressão um tipo de semi-sinônimo<sup>24</sup>; o terceiro termo contempla um termo abrangente, endossado pela TaDiRAH - *Taxonomy of Digital Research Activities in the Humanities*<sup>25</sup>, a principal taxonomia da área, como termo abrangente para todas as práticas das HD. Os termos 3 e 4 garantem a recuperação em línguas relevantes como espanhol e português, muito embora os termos indexadores, bem como resumos, utilizados pelos periódicos e bases de dados são, por regra, em inglês. O termo em francês não necessita ser utilizado, pois a expressão *digital humanities* foi cunhada em inglês na academia francófona, como identifica o Manifesto. O último termo de busca visa contemplar toda a produção que aborda as transformações digitais nas práticas de pesquisa, mas que manejados por atores alheios à expressão oficial e que, por desconhecimento, podem se utilizar da descrição sumária do fenômeno de transformação da metodologia de trabalho pelo universo digital. Dessa maneira a expressão final de busca aplicada às bases pode ser representada da seguinte *string* de pesquisa:

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

Dessa forma o objetivo da revisão foi recuperar itens que tratem da *digital humanities* nas bases de dados consolidadas como grandes repositórios de produção acadêmica. Assim a pesquisa na base Scopus e Web Of Science foi utilizada para refletir um resultado adequado sobre a produção na área. Para tal adotamos como padrão o filtro que contemplasse todas as possibilidades de busca a saber: título, palavras-chave, resumo apenas. Na Scopus essa função atende pelo nome de 'Article title, abstract and keywords' e na Web Of Science de 'Topic'. Somou-se à estratégia de busca a utilização de operadores de lematização<sup>26</sup> que garantem que formas variantes da palavra sejam consideradas na recuperação ou seja, o recuperador entende que a busca deseja as variantes daquela raiz lexical. As aspas respondem na operação de busca pela restrição à expressão grafada, não possibilitando alternativas o que favorece a não recuperação de termos não relacionados. Dessa forma, os termos esforçam-se para estabelecer uma abrangência controlada que nos

---

<sup>24</sup> Ver tópico 1.1

<sup>25</sup> Ver tópico 1.4

<sup>26</sup> A Web Of Science define a lematização como estratégia de redução das formas flexionadas de uma palavra para sua raiz lexical. Com a ativação da lematização, um termo de pesquisa é reduzido ao seu "lema" e formas flexionadas da palavra são recuperadas. Como resultado, a lematização pode reduzir ou eliminar a necessidade do uso de caracteres curinga para recuperar plurais e variações de grafia de uma palavra

permite abarcar não apenas a indexação explícita presente no título e/ou nas palavras-chave, mas a estratégia de considerar o resumo nos viabiliza ampliar o conteúdo potencialmente revelador sobre a temática. A expressão de busca aplicada às bases em 12 de setembro de 2018 recuperou na Web Of Science 1472 documentos e 2483 na Scopus.

#### *4.1.2 Extrair os dados recuperados pela busca: compondo a matriz de análise*

A extração de arquivos de bases de dados tem sido em grande parte o combustível para a escalada dos *softwares* usados para revelar detalhes antes encobertos por um volume memorável de informações. Atualmente há uma alta tendência em analisar dados oriundos de sites, de redes sociais, e também das bases de dados específicas para gerenciamento e recuperação de informações como Web of Science e Scopus, por exemplo. Esse apelo tem se constituído num dos grandes expoentes do movimento das humanidades digitais, tendo seu cerne na aplicação de ferramentas computacionais para análise de grandes volumes de dados. É muito comum que ‘dados’, em tempos de *big data* e explosão informacional, estejam em maior evidência, pois representam a informação em estado genérico e bruto, associados à ideia de volumosos arquivos de informações codificadas, cuja compreensão se dá apenas pelos sistemas de computadores.

Nesse contexto podemos considerar que estão em voga como produto desse processo de produção de informação em grande escala todo tipo de dado codificado. A esses arquivos, denominados tecnicamente como metadados que, de maneira geral, nos dão uma boa noção de tipos de arquivos que podem ser extraídos. Assim, por exemplo, é possível extrair dados de uma plataforma mantida online num servidor de internet a partir de um robô que recolhe essas informações e as entregá-las ao pesquisador em planilhas de dados brutos, melhor aplicáveis a uma determinada análise realizada por máquina. Esse processo denominado "raspagem"<sup>27</sup>, segundo Andriolo (2014), é a atividade de extrair dados de sites e transportá-los para um formato mais simples e maleável que possam ser analisados e cruzados com mais facilidade.

---

<sup>27</sup> Utilizamos o termo técnico da área de computação para acrescentar sentido à ideia de uso tradicional dos resultados obtidos das bases de dados. Assim, consideramos que utilizar dados (ou metadados) que podem ser processados por *software* configura, a nosso ver, uma raspagem de dados haja vista que tradicionalmente apenas os dados quantitativos eram utilizados pelos usuários das bases, sobretudo profissionais da informação. Dessa maneira, para nós, o processo de extração de um arquivo com as informações computacionalmente estruturadas, passíveis portanto de análises mais qualitativas, configura um processo de raspagem.

Nas bases de dados destinadas ao campo da literatura científica, também há a possibilidade de obtenção de dados técnicos que podem ser considerados metadados. É o caso das bases de dados consolidadas em torno do conceito de *Science Citation Index*, que por gerenciar a informação e sua recuperação de forma estruturada, também viabilizam o levantamento da literatura como, por exemplo, as bases Web of Science e a Scopus. Viabilidade essa que explica e justifica sua aplicação nesta pesquisa. O grande potencial dessas bases, por muitas vezes desconhecido, é oferecer formas de exportar os dados cuja leitura se apresenta por meio de formatos tradicionais de consulta, desenvolvidas para operar numa interface amigável ao usuário. Com esses resultados sendo extraídos sob a forma de metadados em arquivos de diversos formatos, como por exemplo, o 'csv' se torna possível o processamento desses dados em *softwares* diversos. Dessa forma ao invés de salvar pesquisas em listas de referências bibliográficas, é possível fazê-lo através em um conjunto de metadados. Assim, no *software* adequado, esses metadados podem revolucionar nossa relação com essas bases e seus conteúdos. Entramos então no terreno da Bibliometria na era das humanidades digitais que habilita o processamento desses recursos através de *softwares* para processar automaticamente métricas como assunto, autoria, área do conhecimento etc. Assim, com a extração de dados e ferramentas computacionais disponíveis ao nosso alcance temos equacionada uma reconfiguração dos métodos tradicionais possibilitado por essa conjugação, que viabiliza um inédito nicho de exploração.

Para Van Eck; Waltman (2014) os pesquisadores da bibliometria têm se interessado por *softwares* voltados à análises bibliométricas por tornar os resultados de suas pesquisas bem mais interessantes devido à associação desses resultados com as ferramentas de visualização, o que tem ampliado largamente o público interessado e o alcance desses estudos. Soma-se ainda a praticidade do automatismo da análise das informações em larga escala. O VOSviewer, por exemplo, é um *software* que atende a essas demandas, viabilizando a construção, análise e visualização de redes de dados bibliográficos. A principal particularidade que nos interessa aqui como diferencial é sua proposta bibliométrica, o que faz desta ferramenta um robusto instrumento para os profissionais da Ciência da Informação. O programa é capaz de construir redes de publicações científicas, revistas científicas, pesquisadores, organizações de pesquisa, países, palavras-chave ou termos. Os itens nestas redes "podem ser conectados por coautoria, coocorrência, citação, acoplamento bibliográfico ou *links* de cocitação" (VAN ECK; WALTMAN 2017, p. 3, tradução nossa).

Assim sendo, nosso objetivo foi aplicar esse programa para extrair da matriz de dados exportada das bases de dados a partir da revisão de literatura, as métricas da bibliometria, além de analisá-las para uma melhor compreensão do domínio estudado e, por fim, apresentar essas métricas em forma de grafos de redes. Dessa forma é possível compreender melhor o domínio através da visualização da complexa interação que os agentes que compõem a estrutura elitista da ciência, responsáveis por formar o Colégio Invisível das humanidades digitais.

#### **4.2 Analisando as métricas e visualizando um campo: aplicando uma estratégia para compreensão do domínio das humanidades digitais**

A aplicação de métodos matemáticos no universo dos livros sempre foi uma operação não integralmente amigável aos profissionais da informação. A constância de fórmulas e equações conferiam aos estudos bibliométricos uma abordagem pouco íntima aos bibliotecários, primeiros responsáveis pela mensuração da produção bibliográfica. Nesse sentido, podemos inferir que as técnicas bibliométricas não se restringiram ao universo da biblioteca, mais que isso, elas foram alvo de aplicação e desenvolvimento por agentes multidisciplinares.

Essa tensão vislumbra no contexto das humanidades digitais mais uma vantagem. Com *softwares* de computador dedicados ao processamento das métricas da bibliometria, grande parte do trabalho complexo foi abreviado. Além disso, o nível de produção do conhecimento elevado a índices exponenciais determinou que práticas como as da bibliometria caberiam também na aplicação de ferramentas informáticas na tentativa de processar mais conhecimento em menos espaço de tempo, aprimorando as técnicas no âmbito deste campo de estudo.

Para Van Eck; Waltman (2014) a ideia de visualizar redes bibliométricas, muitas vezes referida como "mapeamento da ciência", tem recebido séria atenção desde os primórdios da pesquisa bibliométrica. Segundo os autores a visualização revelou-se uma abordagem poderosa para analisar uma grande variedade de redes bibliométricas, "variando de redes de relações de citação entre publicações ou periódicos para redes de relações de coautoria ou redes de relações de coocorrência entre palavras-chave" (VAN ECK; WALTMAN, 2014, p. 1, tradução nossa). Os autores veem o movimento como forma de popularização das práticas bibliométricas tanto no sentido do processamento

dos dados quanto em relação ao produto final dessas análises e, nesse sentido, a estratégia da visualização tem importância ímpar.

Com o tempo, os pesquisadores começaram a analisar redes cada vez maiores, levando à necessidade de técnicas e ferramentas de visualização. Ao mesmo tempo, usuários profissionais de bibliometria, por exemplo, instituições de pesquisa, agências de financiamento e editores, tornaram-se cada vez mais interessados em visualizações de redes bibliométricas. Para disponibilizar visualizações de rede bibliométrica para um público mais amplo, tanto dentro como fora da comunidade de pesquisa bibliométrica, os pesquisadores desenvolveram uma série de ferramentas de *software*. (VAN ECK; WALTMAN, 2014, p. 2, tradução nossa).

Os autores entendem que o cenário de profusão informacional e importância das análises bibliométricas para a própria ciência conduziram esforços no sentido de tornar esse campo de estudo mais difundido e acessível. Todavia não pontuam que esse movimento de dinamização das técnicas aplicadas à bibliometria, bem como as novas formas de se receber seus produtos, configuram-se como mais uma manifestação do fenômeno das *digital humanities*. É nesse contexto que vemos ferramentas aplicadas à essas funções, como VOSviewer, legitimadas e em consonância com as demandas contemporâneas. Para (VAN ECK; WALTMAN, 2014, p. 2, tradução nossa) o VOSviewer é "uma ferramenta que oferece de uma forma relativamente fácil a funcionalidade necessária para visualizar redes bibliométricas". Podemos definir o VOSviewer da seguinte forma

uma ferramenta de software para criar mapas baseados em dados de rede e para visualização e exploração desses mapas. As funcionalidades do VOSviewer podem ser resumidas da seguinte forma: criação de mapas baseados em dados da rede. Um mapa pode ser criado com base em uma rede que já está disponível, mas também é possível construir primeiro essa rede. O VOSviewer pode ser usado para construir redes de publicações científicas, revistas científicas, pesquisadores, organizações de pesquisa, países, palavras-chave ou termos. Os itens dessas redes podem ser conectados por coautoria, coocorrência, citação, acoplamento bibliográfico ou links de citação. [...] O VOSviewer fornece três visualizações de um mapa: a visualização de rede, a visualização de sobreposição e a densidade visualização. A funcionalidade de zoom e rolagem permite que um mapa seja explorado detalhes completos, o que é essencial quando se trabalha com grandes mapas contendo milhares de itens. (VAN ECK; WALTMAN, 2017, p.2, tradução nossa).

Nesse sentido o VOSviewer integrou nossa principal estratégia metodológica para esta pesquisa por nos viabilizar a análise dos pouco mais de 1.836<sup>28</sup> artigos extraídos dessa massa de informações dados como Coautoria e suas respectivas análises de autores, organizações, países; coocorrência sendo possível ainda extrair a relação de palavras-

---

<sup>28</sup> Corresponde ao total de itens após desambiguação

chave; citação de referências, periódicos, autores, organizações e países; acoplamento bibliográfico de referências, periódicos, autores, organizações e países; por fim cocitação de referências citadas, periódicos citados e autores citados.

Apesar de a visualização se aplicar à esta pesquisa enquanto estratégia incisiva, por viabilizar uma topografia concreta do campo, todavia, os grafos são como imagens estáticas de um instante analisado que, de acordo com o movimento natural de crescimento das produções científicas e do aumento do interesse em HDs no mundo, essa imagem - estática - poderá apresentar uma nova configuração, até mesmo antes da conclusão deste estudo. Dessa forma, produzir ciência é cada vez mais compreender-se em um fluxo de produção e comunicação no qual não detemos controle de seu crescimento.

#### 4.2.1 *Análise de citações para o campo das humanidades digitais*

Vimos que a análise de citações compõe um campo importante também da Cientometria, no que pode ser considerado como uma reverberação profunda dos estudos bibliométricos amplamente quantitativos em suas origens. Para LEYDESDORFF (1998, p. 5, tradução nossa) "o impulso primário para organizar as citações em um *Science Citation Index*<sup>29</sup> veio, no entanto, da recuperação da informação". O autor defende que sua existência partiu da conjunção dos interesses em torno do SCI praticado por Eugene Garfield e dos interesses de Solla Price em prover iniciativas de medição desses índices com vistas à recuperação da informação. Para ele esse foi o contexto de "surgimento de conceitos como 'mapeamento da ciência' através da introdução dos estudos da dinâmica das redes de citação" (LEYDESDORFF, 1998, p. 5, tradução nossa). O autor ainda estabelece distinções entre os diversos indicadores bibliométricos, o que o faz definir que as "citações referem-se a outra palavra, frase, período, parágrafo ou título do texto na perspectiva do artigo citante" (LEYDESDORFF, 1998, p. 6, tradução nossa).

Na prática, no entanto, o significado de citações parece mais fácil de entender intuitivamente do que o de códigos analiticamente definidos em textos científicos (Snyder et al., 1995). Além disso, as citações são uma ordem de magnitude mais específica do que, por exemplo, palavras compartilhadas (Leydesdorff 1989). As citações, portanto, parecem ter funções específicas no processo de pesquisa (Garfield, 1996). Neste estudo, argumento que as citações atualmente entendidas surgiram durante uma certa fase do desenvolvimento científico em física, a saber, no início do século XX. **A desconstrução**

---

<sup>29</sup> Ver nota 21

histórica da citação em termos de redes interativas de autores e textos permitirá sugerir uma nova teoria da citação como uma operação dinâmica que permite a redução da complexidade em vários contextos ao mesmo tempo. A perspectiva dinâmica das seleções operando em seleções de outras redes responde pelo caráter das citações como indicadores estatísticos (ou seja, incertos), por sua especificidade e por sua multi-contextualidade. Uma reavaliação das teorias existentes de citação se torna possível, e assim uma abertura para a discussão entre analistas de citação pode ser sugerida. (LEYDESDORFF, 1998, p. 6, tradução nossa, grifo nosso).

Apesar de ter, a nosso ver, debutado os estudos em torno do conceito de mapeamento da ciência, as citações sofreram uma desconstrução com potencial provocado, sugerimos, pela relatividade de sua complexidade de carga cognitiva entre os agentes citante/citados. Esse movimento se fortalece no bojo do crescimento de importância das análises dos demais tipos de análise de citações como as cocitações e acoplamento bibliográfico, por exemplo. No caso da problemática aqui investigada vemos relevância na abordagem das citações, pois essa modalidade de análise está em conformidade, na nossa opinião, no mapeamento de um campo incipiente e complexamente multidisciplinar como o das *digital humanities*.

Nesse sentido, reafirmamos a relevância dessa métrica corroborando com Eugene Garfield (1984) e Loet Leydesdorff (1998) quando observam que as citações, portanto, parecem ter funções específicas no processo de pesquisa. Ou ainda a perspectiva dinâmica "das seleções operando em seleções de outras redes responde pelo caráter das citações como indicadores estatísticos por sua especificidade e por sua multi-contextualidade" (LEYDESDORFF, 1998, p. 6, Tradução nossa). Na nossa avaliação essa qualidade, ainda que meramente estatística, nos agrega à proposta desta pesquisa. Assim sendo, praticamos, a seguir, sua análise no contexto da literatura do campo.

Nos 1.836 documentos recuperados foram citados 4530 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. A escolha do número de autoria ser 3 revela uma tentativa de desconsiderar da análise aqueles autores casuais, que não produzem com recorrência. Por essa razão consideramos que, a partir da produção de 3 artigos, um autor não pode ser considerado casual. A exigência de 6 citações segue a lógica do dobro do número de artigos publicados aplicando mesmo critério de casualidade ao mínimo de citações.



Neste universo o autor que mais produziu artigos possui 15 documentos e o mais citado possui 235 citações. Tanto na produção de artigos quanto na quantidade de citações podemos verificar que os balizadores conceituais das Leis de Zipf e Bradford se confirmam, uma vez que um grupo concentra a maior produção de artigos (6 autores) produz entre 15 e 10 documentos. Em relação às citações essa formação concentradora é ainda mais evidente aonde 7 autores recebem o maior número de citações (de 235 a 110). Para concluir essa discrepância entre peso de produção *versus* quantidade de citações, nos valem do cálculo 'força total do link'<sup>30</sup> equacionado<sup>31</sup> pelo VOSviewer. Considerando esta ordem de grandeza, identificamos que 2 autores se isolam no grupo mais relevante com maior pontuação: um com 35 e o segundo com 24. Um segundo grupo se distribui entre os que alcançaram 18 e o número mínimo de pontuação (1), além de um outro grupo amplo sem nenhum tipo de pontuação no índice 'força total do link'. Podemos representar esse cenário pelo seguinte grafo:

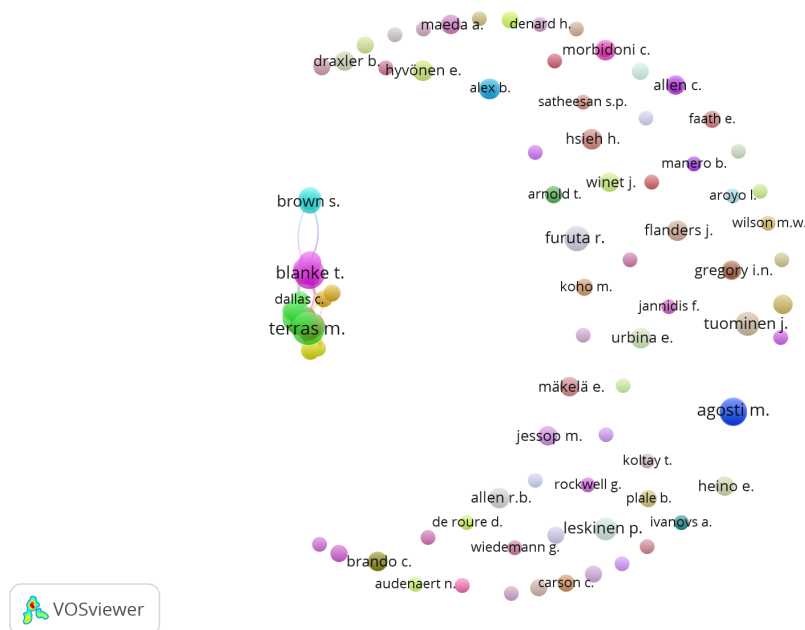


Figura 8: Grafo da rede autores citados (Fonte: VOSviewer).

<sup>30</sup> Total Link Strength: Cada link tem uma força representada por um valor numérico positivo. Quanto maior esse valor, mais forte será o link. A força de um link pode, por exemplo, indicar o número de referências citadas que duas publicações têm em comum (no caso de links de acoplamentos bibliográficos), o número de publicações de dois pesquisadores são coautores (no caso de links de coautoria), ou o número de publicações em que dois termos ocorrem juntos (no caso de coocorrência).

<sup>31</sup> Daqui para frente consideraremos que o índice *total link strength* (força total do link) é a equação representativa que melhor indica os pesos que cada item possui na rede.

O grafo representa com clareza a concentração descrita acima, assim importa ressaltar a característica predominante de pulverização do campo. Foram identificados 76 *clusters*<sup>32</sup> cuja maioria dessas concentrações está isolada e composta por apenas um autor. Contudo podemos visualizar a concentração de 5 clusters (1, vermelho; 2, verde; 5, violeta; 6, azul claro e 8, dourado) composto por 44 nós que mantém *links* mútuos, o que ao nosso ver, revela a estrutura central deste domínio na esfera dos autores citados. Dessa forma podemos sintetizar:

Tabela 2: Autores mais citados por força de link

author	documents	citations	total link strength
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

Fonte: VOSviewer.

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.', apresentando o maior número de citações todavia sem ocupar o primeiro lugar do *ranking*. O VOSviewer calcula a força dos *links* baseados em ligações

<sup>32</sup> Um conjunto de itens junto com os *links* entre os itens correspondem a uma rede. Esses itens podem ser agrupados em *clusters*. Um *cluster* é um conjunto menos (ou mais próximo) de itens incluídos em um mapa.

mútuas, ou seja as ligações estabelecidas por essa autora em específico tem menos relevância que aquelas praticadas por 'blanke t.' e 'hedges m.', por exemplo.

Para compreender os contornos disciplinares das humanidades digitais podemos explorar o contexto acadêmico destes primeiros autores. A área do autor com maior *link* de força (Tobias Blanke<sup>33</sup>) é Filosofia e Ciência da Computação, segundo sua biografia online. Mark Hedges<sup>34</sup>, segundo posicionado, tem formações em Matemática e Filosofia. Stuart Dunn<sup>35</sup> tem formação originária em Arqueologia e está posicionado em terceiro lugar. A autora mais citada de todo levantamento, Melissa Terras<sup>36</sup>, tem origens acadêmicas na História da Arte e Literatura, todavia desenvolve suas atividades no campo da Ciência da Informação, diferentemente dos 3 autores anteriores que atuavam na área de humanidades digitais. Steve Anderson<sup>37</sup> é proveniente da área da História e também atua academicamente neste contexto. A segunda metade dos autores, a exceção de 'brown s.', constituem um *cluster* específico (*cluster* 3, azul escuro), e apesar de relevante, não se integra ao principal agrupamento de *clusters* citados anteriormente. Esse agrupamento evidencia um isolamento dos autores do campo da computação vinculados academicamente a esta área, na qual encontram-se 'Owen Conlan<sup>38</sup>' e 'Cormac Hampson<sup>39</sup>', constatação não diferente de 'Séamus Lawless<sup>40</sup>' e 'Eoin Bailey<sup>41</sup>' que também pertencem ao campo da Ciência da Computação.

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. Todavia, destaca-se que

---

<sup>33</sup> Professor Tobias Blanke. Disponível em: <<https://www.kcl.ac.uk/artshums/depts/ddh/people/academic/blanke/index.aspx>>. Acesso em: 14 set. 2018.

<sup>34</sup> Hedges, Dr Mark. Disponível em: <<https://www.kcl.ac.uk/artshums/depts/ddh/people/academic/hedges/index.aspx>>. Acesso em: 14 set. 2018

<sup>35</sup> About Stuart. Stuart Dunn, 4 abr. 2009. Disponível em: <<https://stuardunn.blog/about/>>. Acesso em: 14 set. 2018

<sup>36</sup> Bio. Melissa Terras, 15 fev. 2017. Disponível em: <<https://melissaterras.org/bio/>>. Acesso em: 14 set. 2018

<sup>37</sup> Steve Anderson Digital. s/d. Disponível em: <<https://steveanderson.digital/cv/>> Acesso em: 14 set. 2018

<sup>38</sup> Owen Conlan. Edge Research, [s.d.]. Disponível em: <<https://edge-research.eu/supervisors/owen-conlan/>>. Acesso em: 15 set. 2018

<sup>39</sup> Cormac Hampson | Trinity College Dublin, Dublin | TCD | School of Computer Science and Statistics. Disponível em: <[https://www.researchgate.net/profile/Cormac\\_Hampson](https://www.researchgate.net/profile/Cormac_Hampson)>. Acesso em: 15 set. 2018.

<sup>40</sup> Séamus Lawless - Assistant Professor. Disponível em: <<https://www.scss.tcd.ie/Seamus.Lawless/index.html>>. Acesso em: 15 set. 2018.

<sup>41</sup> Eoin Bailey | Trinity College Dublin, Dublin | TCD | School of Computer Science and Statistics. Disponível em: <[https://www.researchgate.net/profile/Eoin\\_Bailey](https://www.researchgate.net/profile/Eoin_Bailey)>. Acesso em: 15 set. 2018.

uma autora ligada academicamente ao campo dos estudos informacionais (*Information Studies*) lidere exponencialmente as citações da área. Foi possível identificar também a participação da área da História no levantamento dessa métrica. Além disso, importa constatar a relevância do *cluster* de autores da Ciência da Computação, bem como frisar seu isolamento. Assim, a participação de autores que não formalmente ligados, em termos institucionais às *digital humanities*, são Ciência da Informação e Computação.

#### 4.2.2 Análise das fontes de publicação para o campo das humanidades digitais

Os periódicos científicos cumprem um papel fundamental no desenvolvimento da ciência por sua função de promover a divulgação científica. A relevância da função cumprida por esses periódicos é reconhecida pela bibliometria, estando, nesse escopo, também sujeitos à análises tanto quanto as citações. Nesse sentido vemos na Lei de Bradford a expressão dessa importância, que propôs em 1948, a formulação de sua lei para compreensão da distribuição dos periódicos

se as revistas científicas forem organizadas em ordem decrescente de produtividade de artigos sobre um determinado assunto, elas podem ser divididas em um núcleo de periódicos mais particularmente dedicados ao assunto e vários grupos ou zonas contendo o mesmo número de artigos que o número de periódicos no núcleo". (BROOKES, 1969, p. 953, tradução nossa).

A formulação teórica para essa lei segue, a nosso ver, no sentido também da lógica de concentração da ciência por equacionar a indicação de que um grupo restrito concentrará a maior parte da produção de determinada área ou temática. Sintetiza COUTINHO (1991, p.170) que Bradford “observou o comportamento repetitivo da literatura técnica sobre determinado assunto: poucos periódicos produziam muitos artigos e muitos produziam poucos, havendo ainda dispersão da literatura”, isto é, grande número de artigos sobre determinado assunto divulgados em periódicos devotados a assuntos diferentes.

Em nossa análise observamos um total de 1154 fontes e para estratégia de refino consideramos que o periódico deveria deter a qualidade de 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. O periódico mais ativo foi da área de Ciência da Computação, publicou

96 artigos, estando em destaque no topo deste ranqueamento (em segundo lugar estão empatados 2 periódicos com 77 publicações). O mais citado, da área de Computação e Linguística, se estabelece como o mais citado com 326 citações (o segundo mais citado apresenta 285 citações).

Quando ranqueados por 'força de *links*' a disposição dos títulos se altera consideravelmente. Um grupo de 3 periódicos se isola no topo com força de *link* entre 53 e 50; um segundo grupo de 10 títulos se distribui entre índices 39 e 10; e os demais com pouca relevância (26 títulos) se dispersam com pouca literatura (índice de 8 a 1). Ainda assim um grupo amplo se apresenta com relevância zero para avaliação de rede de estabelecimento de *links*. Dessa forma podemos representar essas condições com o seguinte grafo:

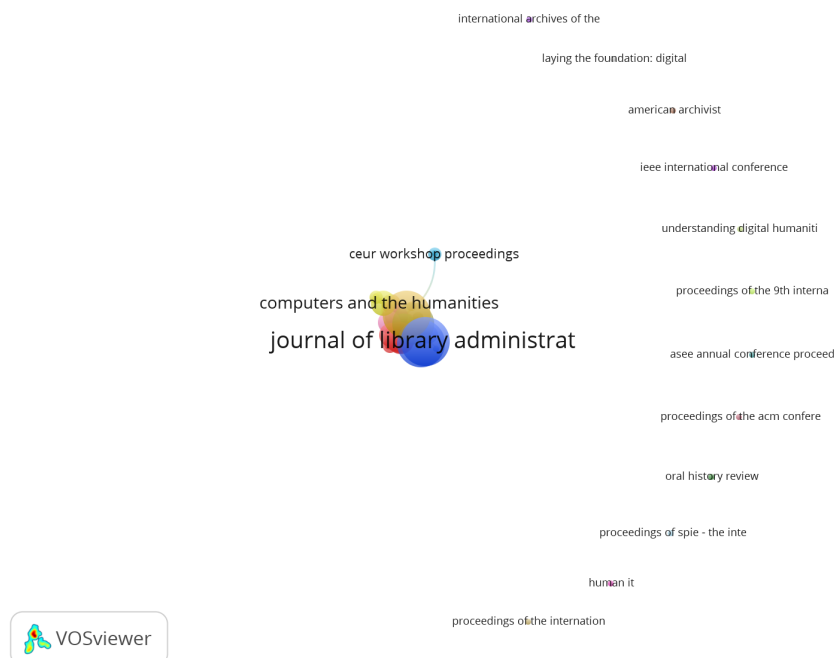


Figura 9: Grafo periódicos mais citados (Fonte: VOSviewer).

O grafo segue corroborando a noção fragmentada do campo das humanidades digitais, todavia segue em consonância com a teorização da Lei de Bradford. Do total de 51 títulos, 12 podem ser desconsiderados por não praticarem *links* relevantes, permanecendo, portanto, isolados. Há um núcleo concentrando no estabelecimento de *links*: são 39 formando esta condição. O grafo compõem-se de 22 *clusters*, sendo destes 12 compostos por um título e, por isso, isolados sem estabelecimento de ligações. Os *clusters* conectados mutuamente totalizam 10 e não é possível identificar uma área

predominante para a maioria deles, a exceção do *cluster* 4 (amarelo) que integra-se por 4 títulos, sendo 3 devotados à área de bibliotecas. Dessa maneira podemos sintetizar na seguinte tabela:

Tabela 3: Periódicos mais citados por força de *links*.

source	documents	citations	total link
			trengh
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

Fonte: VOSviewer.

A tabela esclarece 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. Neste caso podendo ser entendido como aquele com maior aproveitamento de suas publicações já que foi este que possui mais citações em relação ao número de artigos publicados. Ainda destaca-se essa área como a de maior contundência nesse ranqueamento por apresentar 3 títulos, além da primeira posição, também a terceira e quarta colocações. Apesar dos periódicos desta lista

revelarem uma pulverização multidisciplinar, pode-se identificar ainda a presença das áreas da Computação, Linguística e História.

A análise desta métrica 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. Todavia é importante ponderar que a presença de títulos (no caso há apenas um nessa condição) produto de eventos científicos como o caso '*conference proceedings*' pode contribuir com a dispersão de artigos, descentralizando em relação a uma área. Assim sendo, a dispersão disciplinar é corroborada, apesar de ser possível identificar algumas áreas com presença declarada e apenas uma com participação destacada.

#### 4.2.3 Análise dos documentos mais citados na literatura das humanidades digitais

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 sugiram indicações cognitivas. Assim pretendemos submeter os mais de 1800 itens recuperados sob análise de citações para identificar quais destes documentos concentram relevância enquanto influência intelectual nos artigos que o citaram. Do universo recuperado, 1231 sequer receberam qualquer citação, e 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. Contudo, esses documentos tem relevância mínima ou nula quando observada criação de *links*. Como a relevância atribuída pela nossa análise centra-se na capacidade de estabelecimento de ligações, passamos a analisar a distribuição segundo esse ordenamento.

Apenas 4 documentos se estabelecem enquanto grupo com maior constituição de *links* no universo de mais citados, a distribuição se apresenta com 2 documentos com constituição de 3 *links* e, outros 2, com 2 *links*. Na sequência 14 documentos formam apenas 1 *link*, a grande maioria dos documentos não formam nenhum *link*. Esse cenário pode ser percebido na apresentação do grafo:

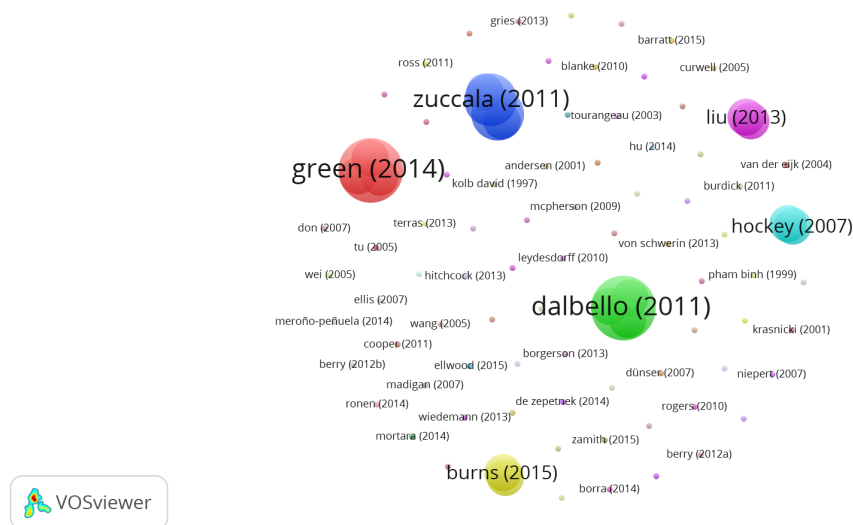


Figura 10: Grafo de documentos com maior número de links entre os mais citados.

O grafo destaca os documentos mais relevantes para essa análise, indicando-os através dos 'nós' mais proeminentes, além disso deve-se pontuar o fato desses itens se destacarem em relação a um todo com um índice de dispersão acentuado. Esse quesito atribui importância central aos nós em destaque, pois além de enquanto citados estabelecerem *links* (recebem ou promovem citações) com outros documentos também relevantes. Essa rede é composta por 81 *clusters* dos quais 75 são compostos por apenas um item e não promovem ligações com nenhum outro nó. Dos 6 *clusters* que possuem relevância, a primeira metade possui 4 documentos (*clusters* 1, 2 e 3), enquanto os demais, 2. Uma observação que se destaca nesta análise é o isolamento dos *clusters* mais relevantes, até agora a tendência foi de que *clusters* relevantes exerciam ligações mútuas com outros, nesse recorte essa tendência não se confirma.

Podemos perceber ainda, no que tange a moldura intelectual dos *clusters* concedida pela proveniência dos documentos, que os agrupamentos 1 (vermelho) e 3 (azul escuro) são compostos por itens publicados em periódicos da área de biblioteca (*Information Studies*). O *cluster* 2 (verde) também possui metade de seus itens proveniente da mesma área dos anteriores. Apenas o *cluster* 4 (amarelo) possui determinação intelectual perceptível em sua composição, tendo todos os 2 documentos proveniente da Geografia, itens dessa área ainda integram, isoladamente, outro *cluster* (2 verde). Aparecem ainda alguns documentos proveniente da Computação (1 no *cluster* 2,



verde; e 1 no *cluster* 6, azul claro). Em síntese, podemos elencar os 10 mais relevantes documentos citados segundo a análise:

Tabela 4: Documentos mais citados por força de links

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

Fonte: VOSviewer.

Os quatro documentos mais relevantes desta métrica 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)*. As posições 5<sup>a</sup>, 6<sup>a</sup> e 8<sup>a</sup> são ocupadas por documentos publicados pela área da Geografia (*Geojournal*, *Geoforum e Progress in Human Geography*, respectivamente). Aparece ainda publicação da área de Linguística e o último documento também é proveniente da área de biblioteca (*Journal of Library Administration*).

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. 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. Vemos 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.

### 4.3 Análise de coautoria entre autores

O VOSviewer é uma ferramenta de aplicabilidade ímpar para análises bibliométricas, permitindo dentre as diversas possibilidades de observação, desde análise de métricas mais tradicionais às menos usuais. O exemplo disso é sua capacidade de cruzar dados de coautoria não apenas de autores – uma análise mais usual, mas também entre Organizações nas quais esses autores são afiliados e entre países de acordo com a nacionalidade dos mesmos. Para fins de análise neste tópico, nos dedicaremos aqui ao usual, que versa sobre a coautoria de autores, reservando as outras duas possibilidades para um tópico específico sobre fronteiras e barreiras geográficas das humanidades digitais.

A relevância da colaboração científica expressada pela coautoria, sempre trouxe consigo relevância ímpar para os estudos de análise de disseminação da ciência e, nesse sentido revela-se, não apenas na colaboração de característica bibliométrica, mas também em contornos não bibliométricos como defendem Ponomariov; Boardman (2016). Para estes autores existe uma importante questão na forma de colaboração que nem sempre segue expressa por elementos bibliométricos. Apesar de reconhecerem que estudos baseados em objetivos de mensuração de colaboração não expressa bibliometricamente, são nestes estudos “- usando dados bibliométricos – que modelam a coautoria como resultado de algum tipo de tratamento, como uma nova política ou esses estudos a usam para descrever diferenças nos padrões de publicação entre grupos, para instâncias entre disciplinas”. (PONOMARIOV; BOARDMAN, 2016, p. 1941-1942, tradução nossa). Esse é um ponto que nos importa neste estudo: observar a estrutura formal da participação de autores na produção da literatura das HDs a despeito das outras formas de colaboração entre autores e a problemática que a cerca.

Dessa maneira nos importa focar essa colaboração enquanto uma engrenagem que ilustre as redes de construção do campo a partir da associação de indivíduos em torno de suas produções acadêmicas. Essas associações correspondem a uma estrutura que pode ser vislumbrada enquanto redes colaborativas cujo

conceito de rede é aplicado em diversas áreas do conhecimento humano. Genericamente, pode-se definir uma rede como um conjunto de elementos que mantêm conexões uns com os outros. Na literatura matemática, as redes são reconhecidas como grafos, seus elementos como vértices e suas conexões como arestas. Já nas ciências sociais, os elementos são denominados atores e suas conexões são laços. Já na literatura da ciência da computação, os elementos são reconhecidos como nós e as conexões como ligações. (BRANDÃO; PARREIRAS; SILVA, 2007, p.110-111).

Assim a complementação da análise da métrica coautoria utilizando um *software* nos é central, pois possibilita produzir uma visualização da rede de autoridades associadas centrais no espectro das humanidades digitais, sendo úteis essencialmente na reverberação dos núcleos das distintas áreas acadêmicas que se revelam em proeminência. Dessa maneira prosseguimos com a análise na métrica de coautoria neste estudo bibliométrico como forma de alcançar os atores e os laços de suas principais conexões.

Para análise desta categoria processamos via software 3453 autores na matriz construída e, para estes, estabelecemos como linha de corte os autores que possuísem no mínimo 5 documentos publicados. Essa exigência segue a ideia da criação de parâmetros das métricas anteriores, na qual nos preocupamos em exigir um número pelo menos maior que 3 publicações assinadas para diminuir o risco da incidência de autores esporádicos, condição que tem nas humanidades digitais uma acentuação na banalidade da colaboração acadêmica.

Somamos a essa exigência a necessidade de ser citado pelo menos iguais 5 vezes e, nessas condições, atingimos o universo de 31 autores mais relevantes na análise de coautoria que estão capacitados à análise por possuírem, no mínimo, 5 autorias e citações em documentos com coautoria. Assim temos no grafo abaixo as redes de coautoria nas humanidades digitais:

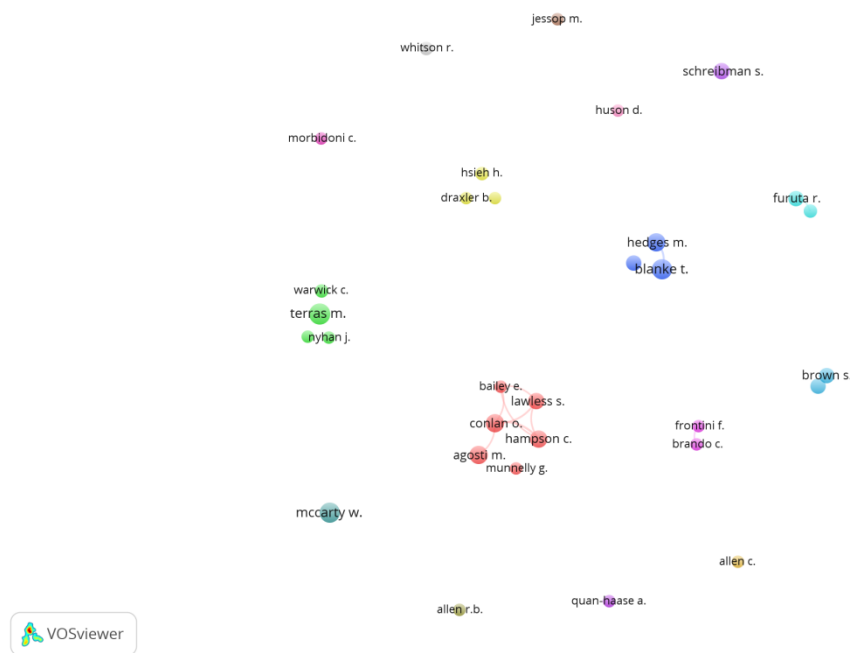


Figura 11: Grafo de coautoria de autores.

O grafo segue comprovando a setorização dos núcleos de 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. Dessa maneira pode-se constatar 4 *clusters* predominantes que variam entre 6 'nós' o maior na cor vermelha e o menor com 3 'nós' na cor amarela. Todavia é importante que se ressalte que o *link* de força aqui se estabelece pela quantidade de coautores versus coautores que fazem parte do recorte. Devemos pontuar essa particularidade, pois o segundo *cluster* possui autores com maior produtividade e igual capilaridade de suas publicações. Ou seja, o agrupamento vermelho possui mais membros em coautoria, todavia existe um tipo de reciprocidade entre seus membros (no momento desta análise praticamente todos os autores do *cluster* tem a mesma afiliação institucional) que o qualifica enquanto maior *link* de força. Enquanto isso o segundo *cluster* (de cor verde) apresenta maior capilaridade, qualidade essa presente na quantidade de produção de documentos e de citação.

Abaixo apresentamos a tabela com o ranqueamento dos autores que nos permite identificar a contraposição ao cálculo do *link* de força que atribui ao *cluster* que não tem a liderança na produção dos documentos nem na quantidade de citação.

Tabela 5: Todos os autores em coautoria por força de links.

author	documents	citations	total link strength
conlan o.	10	44	31
hampson c.	9	43	29
lawless s.	8	31	23
bailey e.	5	29	18
munnelly g.	5	18	17
agosti m.	10	33	16
blanke t.	13	95	12
hedges m.	10	66	11
hsieh h.	6	6	10
draxler b.	5	6	10
winet j.	5	6	10
terras m.	14	200	7
dunn s.	7	40	7
furuta r.	7	35	6
urbina e.	6	18	6
brando c.	5	9	5
frontini f.	5	9	5
warwick c.	6	174	3
nyhan j.	5	56	3
vanhoutte e.	5	49	3
hockey s.	7	61	2
brown s.	7	25	2

Fonte: VOSviewer.

Escolhemos elencar nesta tabela a totalidade dos autores com expressão, segundo sua força de *link*, para ilustrar que o algoritmo do *software* privilegia, enquanto *cluster* principal, aquele que possui o maior número de autores em coautoria. A despeito disso, colocamos foco no segundo *cluster*, pois este possui autores com quantidade extremamente maior de artigos assinados e também, principalmente, de citações. Não discordando do cálculo do algoritmo –que atribui relevância aos mais destacados tendo a força do *link* sendo provida por essa relação: destacados em relação direta com outros também destacados-, mas confrontando-o com o fato de que, apesar de relevância secundária, a problematização do *cluster* 2 nos é pertinente. O *cluster* 2 (de cor verde)

possui a autora com o maior número de artigos produzidos e, talvez por isso, também mais citada. Seguido a ela também há outros 3 autores possuem números altamente relevantes para esta análise, uma vez que demonstram uma capacidade maior de penetração. Essa capacidade é possibilitada pela capilaridade com a qual indivíduos da área de *Information Studies*<sup>42</sup> se relacionam com outros autores em coautoria, ou seja a coautoria dos membros deste *cluster* não está concentrada, diferentemente dos membros do *cluster* principal. Ainda assim podemos ter no alto índice de citações a comprovação da capacidade de pulverização dessas publicações em áreas não concentradas.

De fato a análise de coautoria reflete, naturalmente, a relevância de obras referenciais para uma área recente 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* tem maior influência do que os atores do *cluster* principal (vermelho) por aqueles que promovem coautoria com um número significativamente maior de autores e principalmente por serem citados massivamente por, como inferimos, indivíduos de diversas áreas do conhecimento. Assim, defendemos que os autores da área de *Information Studies* tem maior relevância no papel delineador do campo considerando a análise de coautoria.

#### **4.4 Estruturando as cocitações para contribuição da percepção do campo das humanidades digitais**

Uma métrica de peculiar relevância nos estudos bibliométricos é o estudo das cocitações. Sua percepção remete a uma sensibilidade sobre determinado domínio, pois nas ações de cocitação se revelam detalhes de 2 autores e/ou obras sendo citados em conjunto por uma mesma obra e/ou autor. Nesse sentido

Os estudos de análise de cocitação, baseados na frequência com que dois autores ou documentos são citados de forma conjunta na produção científica de uma área, evidenciam como a estrutura de conhecimento de uma área é percebida pelos pesquisadores. Tem como princípio o fato de que, quando dois documentos ou autores são citados juntos em um trabalho posterior, existe, na perspectiva do autor citante, uma proximidade de assunto entre os citados. Assim, quanto maior a frequência de cocitação, mais próxima a relação entre esses autores citados. (GRÁCIO; OLIVEIRA, 2016, p.197).

---

<sup>42</sup> Ver nota 31 para TERRAS, Melissa; WARWIC, Claire (ver <https://www.dur.ac.uk/about/governance/senior/cwarwick/>); NYHAN, Julianne (ver <https://www.ucl.ac.uk/information-studies/julianne-nyhan>) e VANHOUTTE, Edward (ver <http://ucl.academia.edu/EdwardVanhoutte>) que atuam academicamente na área de *Information Studies*.

Como pontuam as autoras, o principal atributo da cocitação é externar a proximidade entre autores. Essa proximidade acaba por ser responsável por uma tecitura de uma área, de um campo ou ainda de uma disciplina. Nosso ponto é que esta qualidade da cocitação, quando analisada sobre um universo relativamente pequeno, potencializa-se e agrega ainda mais relevância à interpretação do domínio de uma área/disciplina como o caso das humanidades digitais.

Apesar de existirem alguns tipos de indicadores matemáticos<sup>43</sup> capazes de atribuir grandeza numérica aos indicadores de proximidade, defendemos que a opção de simular essa proximidade com grafos se rende a um apelo à visualidade que promove mais atratividade à sua interpretação. Como a cocitação é caracterizada pela citação de ambos itens por um mesmo indivíduo, creditamos ao grafo de rede a simplificação em exibir esta relação, criando maneiras de promover visualmente o que classificamos enquanto tecitura do campo.

Dessa forma a importância de analisar a frequência em que documentos citados estão juntos em um artigo se apresenta, segundo OLIVEIRA; ALVES (2017, p.236) "para a compreensão de um contexto e das relações que acontecem em uma comunidade científica". Assim a tecitura a qual referimos pode ser compreendida pelo papel que a cocitação exerce em uma área pois,

O número de vezes que os autores ou documentos foram citados juntos determina a força da cocitação entre eles, bem como o núcleo da literatura da área. As citações indicam os paradigmas das comunidades formadas, seus procedimentos metodológicos, os grupos de cientistas, suas publicações, e evidenciam os pesquisadores de impacto de uma área. (OLIVEIRA; ALVES, 2017, p.236).

Com efeito, se na análise de citação vislumbramos a grandeza de um item ou indivíduo, na cocitação essa medida se fortalece com a presença de mais um item citado em conjunto. Como a área de um autor pode ser reveladora para os estudos do domínio, podemos ter na reincidência de uma mesma área de dois autores ainda mais relevância

---

43 São exemplos o Cosseno de Salton, o Índice de Jaccard e a Correlação de Pearson que se aplicam a níveis distintos de estudos bibliométricos. Segundo OLIVEIRA; ALVES (2017, p.235-236) esses níveis podem ser entendidos enquanto "Bibliometria para profissionais da bibliometria; Bibliometria aplicada às disciplinas científicas; Bibliometria para a política científica e gestão." Para os autores a aplicação desses indicadores servem principalmente ao primeiro grupo nos quais "são próprio da pesquisa bibliométrica "de base", que está preocupada com o seu desenvolvimento conceitual-teórico-metodológico"

para compreensão do domínio, sobretudo das humanidades digitais. Por certo esta é uma análise altamente pertinente para a *digital humanities* por se tratar esta de um campo incipiente que carece de compreensão dos paradigmas de suas comunidades integradas, seus grupos de cientistas, além de seus pesquisadores de impacto.

#### 4.4.1 A cocitação entre autores

Na cocitação podemos analisar, utilizando a metodologia do VOSviewer, a métrica por três pontos distintos a saber: de autores, de referências e de fontes. Assim utilizaremos a métrica 'por autores' e complementaremos essa análise de cocitações também apresentando os dados respectivos às referências e fontes.

A matriz dos dados possui 38.617 autores 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.

Dessa maneira alcançamos o seguinte grafo de cocitação de autores:

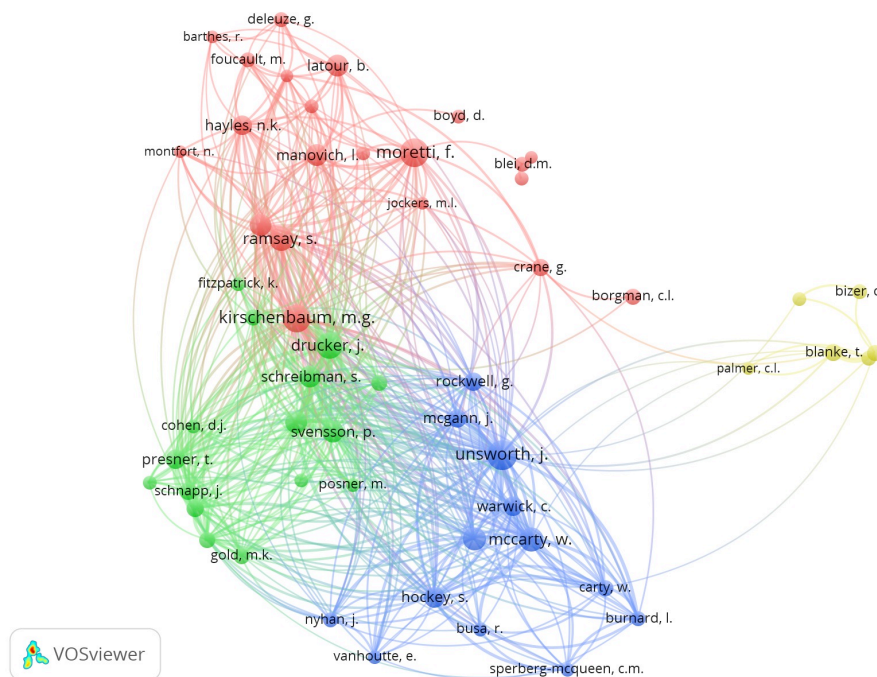


Figura 12: Grafo de cocitação de autores.



O grafo acima constitui-se de 5 *clusters*, dos quais o maior e mais representativo (de cor vermelha) possui 21 autores, seguido com certo equilíbrio, pelo segundo *cluster* (de cor verde) com 16 autores. O *cluster* 3 (azul) é composto por 13 autores e o quarto (amarelo), que juntamente com o quinto, são aqueles que apresentam menor integração com os 3 principais. O último, com apenas dois indivíduos, fora retirado da ilustração para não prejudicar visualização por estar completamente afastado dos demais.

O grafo de cocitação apresenta uma importante inovação em relação aos observados até agora, neste há uma integração entre *clusters* muito evidente, fator completamente divergente das análises de citação e coautoria, por exemplo. Ao nosso ver, esse comportamento demonstra com evidência que os grupos acadêmicos ora distantes, seguem em aproximação, sugerindo um movimento de maturação - pelo menos teórico-literária - do campo. Ao citar indivíduos de áreas distintas, o campo começa a perder a identidade pluridisciplinar e a construir uma identidade própria do "novo" campo. Se em citação e coautoria a tendência era utilizar pares acadêmicos de suas próprias áreas de atuação, agora nesse cenário esboça-se diversificar, mostrando que indivíduos de áreas que não as mesmas começam a citar outros que não de suas próprias áreas. Assim, supomos, há por parte dos autores a busca por outros especialistas reconhecidos com referência na reflexão das *digital humanities*, o que acaba por criar autores referenciais além do reduto acadêmico daquele autor e essa métrica pode ser responsável por permitir a identificação desse fenômeno.

Uma outra prova de que esse cenário se confirma é observar os indivíduos mais citados em cocitação. Quando colocamos esse dado à prova percebemos nitidamente que, diferentemente dos cenários anteriores de citação e coautoria, não há uma predominância de autores de uma mesma área nos primeiros ranqueados. Pela primeira vez não é possível apontar que um grupo de determinada área acadêmica estabelece-se frente outra. Dessa forma temos

Tabela 6: Autores mais cocitados por número de citações.

author	total link	
	citations	strength
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

Fonte: VOSviewer.

A tabela acima 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.'. Na nossa leitura temos nesta métrica a evidência de que autores como os dois mencionados atuam como expoentes epistemológicos do campo, sobretudo se recapitularmos que 'mccarty, w.', por exemplo, é dos expoentes das reflexões sobre '*Humanities Computing*', precursora das humanidades digitais.

Por outro lado destacamos a constatação de novos autores despontarem no ranqueamento por cocitação, como é o caso de 'kirschenbaum, m.g.'. Nome 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 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 o caso de 'terras, m.', 'unsworth, j.', 'drucker, j.' e 'ramsay, s.'. Assim, consideramos pertinente fornecer o grafo de densidade dessa métrica para auxiliar a compreensão do comportamento apresentado pelos autores nessa análise:

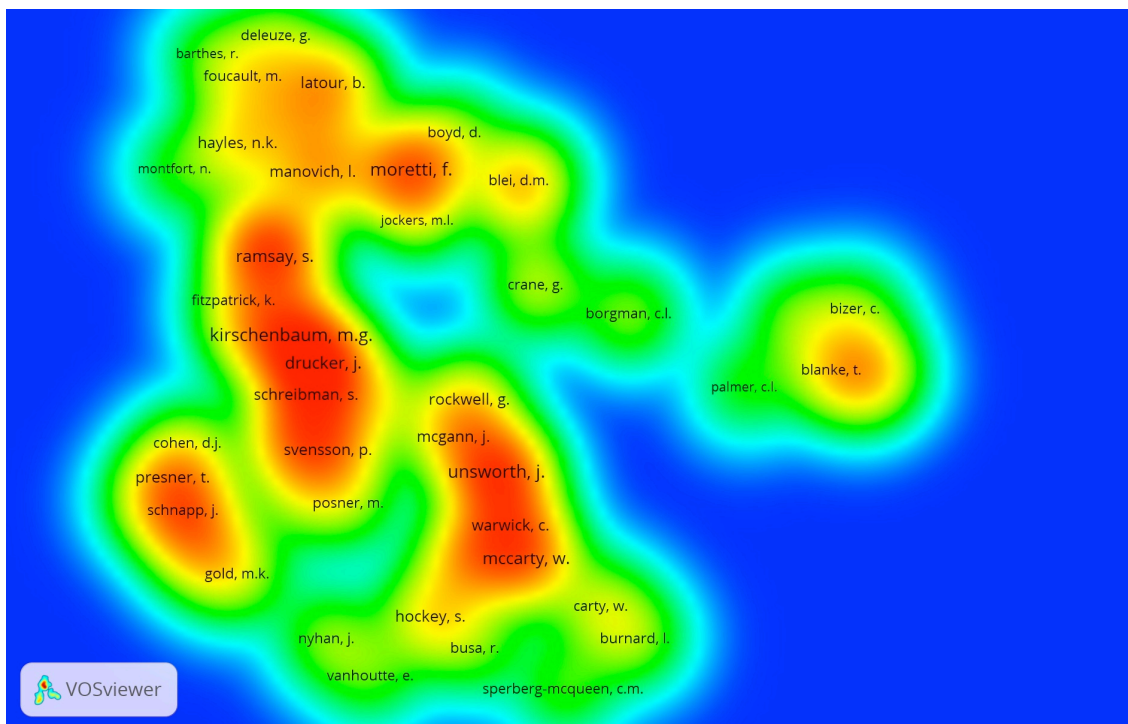


Figura 13: Grafo de densidade de cocitação de autores.

O grafo de densidade nos fornece uma imagem em que a cor mais vibrante (vermelho) indica alta concentração de importância desses indivíduos, ajudando na compreensão de que aqueles 'nós' das áreas vibrantes são os mais destacados quando por ocasião na análise praticada.

#### 4.4.2 Referências de destaque nas cocitações: publicações mais influentes nas humanidades digitais

Se a praxe do uso dos autores para análise de cocitação é uma constante, consideramos válido complementar essa métrica analisando também as referências em destaque quando cocitadas. Ao passo que essas referências oferecem detalhes importantes, nos valemos dessa riqueza de informações para identificar a proveniência desses artigos como forma de obter subsídios para julgar a influência de suas áreas acadêmicas de origem no domínio das humanidades digitais.

Assim sendo de 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 alta, é preciso garantir que nenhuma das

referências que compõem o grafo integre o mesmo por acaso. Assim chegamos ao número de 11 artigos mais cocitados e cujo grafo exibimos, em perspectiva, abaixo:



Figura 14: Grafo de cocitação de referências - cluster 2 e 3.

O mapa de rede das referências cocitadas constitui 3 *clusters*, sendo que apenas 2 estão em conexão. O primeiro (não exibido no grafo acima) encontra-se isolado e compõe-se de 5 referências. O segundo é formado por 4 referências cuja temática versa objetivamente sobre bibliotecas. Importa destacar que a conexão entre os *clusters* 2 e 3 conjuga cocitações entre obras centrais para área (cor azul) com outras que abordam temáticas de bibliotecas (cor verde). Esse detalhamento pode ser conferido na tabela abaixo que lista todas as 11 referências mais cocitadas.

Tabela 7: Referências cocitadas por força de links.

cited reference	total	
	link	strength
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
blei, d.m., ng, a.y., jordan, m.i., latent dirichlet allocation (2003) journal of machine learning research, 3, pp. 993-1022	7	1

Fonte: VOSviewer.

Podemos verificar na tabela que apresenta as referências, que o *cluster 2* é, em totalidade, o responsável por deter a maior força de *link*, ou seja, todos seus itens possuem as maiores relevâncias nas citações. 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.

Paralelamente o *cluster* maior (vermelho) é formado por artigos que tratam de temas não contundentes, todavia que versam sobre a temática da tecnologia. Apesar de não constituírem referências centrais, apresenta o artigo de 'busa, r.' que configurou-se enquanto o principal marco temporal das reflexões sobre *digital humanities*. Esse artigo tem posição central na epistême do campo por abordar aquela que é considerada a primeira ação de *Humanities Computing* através da compilação, em meio eletrônico, da obra de São Thomás de Aquino pelo Padre Jesuíta Roberto Busa: *The Index Thomisticus*. Mesmo sendo considerada como parte da genealogia do campo, a obra não circula entre as cinco principais cocitações o que, ao nosso ver, indica que as reflexões sobre surgimento das humanidades digitais, bem como sua afirmação enquanto campo, têm perdido escala. Com essa perda de escala podemos inferir que os humanistas digitais deixam, paulatinamente, de fundamentar epistemicamente a área dissertando menos sobre seu surgimento. Em paralelo, com o notório crescimento da literatura do campo, podemos supor que a área se encontra afirmada em termos de estabelecimento acadêmico e em pleno crescimento reflexivo.

A métrica de cocitações, sobretudo esta que foca nas referências, revela detalhes importantes sobre a campo através de sua literatura. Com o rigor de se extrair dentre milhares de referências aquelas de maior peso na tecitura do campo nos permite supor que, apesar de toda a característica pluridisciplinar das humanidades digitais e seu apelo tecnológico, o campo da informação assume um protagonismo ímpar por ditar com a temática das bibliotecas as principais problematizações no domínio das *digital humanities*.

#### 4.4.3 Complementando a relevância das referências cocitadas: analisando a cocitação de fontes

Com a cocitação de referências concedendo um peso importante nas reflexões sobre o campo e seu domínio, com uma análise sobre uma métrica também respectiva ao teor bibliográfico da literatura, podemos complementar essa complexa análise. Quando nos referimos às cocitações de fontes, abordamos os veículos pelos quais se vetorizam as

referências, ou seja consideramos nesta métrica os títulos que hospedam as referências. Dessa forma, enquanto fontes, podem ser consideradas para fins de análise, periódicos, livros, capítulos ou qualquer publicação que possua referenciamento bibliográfico. Por essa razão e, não menos por considerar o papel da análise dos periódicos para os estudos bibliométricos, nos debruçamos sobre aqueles que são os principais periódicos cocitados no objeto dos dados aqui construídos.

Nosso corpus de estudo possui 28.740 fontes e para possibilitar a análise, 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. Com essas exigências alcançamos um grafo compreensível com conexões entre todos os *clusters* a conferir:

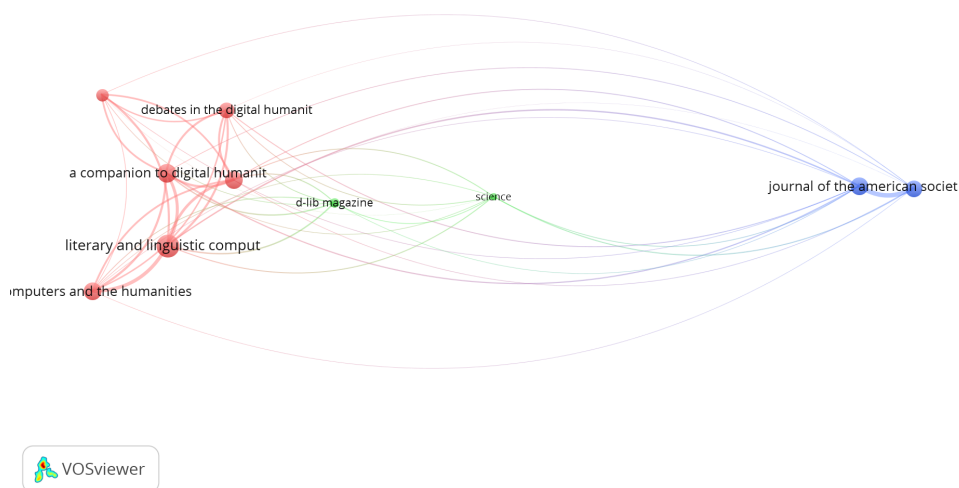


Figura 15: Grafo de cocitação de fontes.

Esta rede forma-se por 3 *clusters* no total, estando os três conectados entre si, reafirmando, ao nosso ver, que uma perda de escala vem se dando na máxima da fragmentação pluridisciplinar do campo. O primeiro *cluster* (cor vermelha) pode ser considerado o núcleo da rede, haja vista que possui 6 itens, supondo enquanto acessório o papel dos outros dois. Os *clusters* 2 (cor verde) e 3 (azul) são compostos por 2 itens cada.

O *cluster* central (número 1, de cor vermelha) representa com significância a composição do campo considerando as áreas acadêmicas mais proeminentes que verificamos até aqui. Nele estão presentes duas obras centrais para área ('*debates in the digital humanities*' e '*a companion to digital humanities*'). Essas obras podem ser consideradas como responsáveis por cunhar o campo através de seu rótulo *digital*

*humanities* o que explicaria sua importância e consequente presença de destaque nessa análise. Além desses títulos icônicos, também merece ser destacada a presença diminuta dos periódicos da área da computação cuja representatividade se baseia apenas com o '*computers and the humanities*'. Merece destaque que um único periódico da Literatura/Linguística (*literary and linguistic computing*<sup>44</sup>) tenha sido o mais citado em citações. Por outro lado a área mostra-se cada vez mais autoafirmada através da literatura, como na relevância do orgânico '*digital humanities quarterly*' que veicula-se como o segundo maior citado nas citações.

O *cluster* central também reúne periódicos da área informacional como por exemplo '*journal of library administration*' periódico responsável por publicar todas as 4 referências mais citadas (como analisado no item 4.4.2). Com essa recorrência a temática de domínio deste periódico (bibliotecas) e sua respectiva área (de *Information Science/Library Studies*) mantém-se presente em mais uma métrica representativa para a tecitura das humanidades digitais.

Tabela 8: Principais fontes citadas.

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

Fonte: VOSviewer.

Não obstante temos ainda a área respectiva aos estudos informacionais presente em todos os outros dois *clusters*: única área acadêmica a possuir essa presença (*d-lib*

<sup>44</sup> Atualmente denomina-se *Digital Scholarship in the Humanities*



*magazine'* no *cluster 2* e '*journal of the american society for information science and technology'* no *cluster 3*). O grafo 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.

Com efeito a cocitação de fontes incrementa nossa compreensão pela análise dos dados das referências, além das relações entre autorias, complementando uma interpretação sobre os dados bibliográficos. Possibilita ainda vislumbrar os caminhos pelos quais autores veiculam suas publicações criando nessa condição uma maneira de clarificarmos quais áreas estão participando também do crescimento da literatura do campo. Ainda assim esboçam-se através desses periódicos em análise aqueles que vem sendo potencialmente apontados como referências para as publicações, desvelando conseqüentemente suas áreas de domínio com destaques indeclináveis para as humanidades digitais.

A análise de cocitação é um método único para estudar a estrutura cognitiva da ciência. A análise de cocitação envolve o acompanhamento de pares de artigos que são citados juntos nos artigos de origem. Quando os mesmos pares de artigos são cocitados por muitos autores, grupos de pesquisa começam a se formar. Os artigos cocitados nesses grupos tendem a compartilhar um tema comum. Combinado com técnicas de agrupamento e dimensionamento multidimensional, a análise de cocitação pode literalmente mapear a estrutura de áreas de pesquisa especializadas, bem como a ciência como um todo. (SURWASE, G., et al. 2011, p.180-181).

Por diversas razões essa análise aplica-se, como vimos, adequadamente às humanidades digitais por revelar cenários correntes, todavia não tão compreendidos até então. Para os autores esse tipo de análise encoraja pesquisadores a explorar a estrutura intelectual da ciência, além de orientá-los sobre os seus desenvolvimentos futuros. Um dos recursos mais valiosos no caso da aplicação neste contexto trata-se da revelação das ligações intelectuais estabelecidas através do processo de referência e citação. É uma ferramenta de pesquisa para examinar o intelectual desenvolvimento e estrutura da disciplina científica. Além disso o fato de que "grupos tendem a compartilhar um tema comum" muito se insere nas práticas acadêmicas das humanidades digitais sugerindo potencialmente a formação do campo pelo que o autor indica enquanto a citação recorrente de pares de autores ser responsável pela formação de grupos de pesquisa. Este

é, senão um proeminente cenário das *digital humanities*, seu diagnóstico mais tangível até então.

#### 4.5 O papel do acoplamento bibliográfico para o domínio

Na esteira da influência que um artigo possui ao citar outras duas referências, também se promoveu a relevância advinda do fato de que dois artigos têm ao citar um ou mais artigos idênticos. A potencial relação de proximidade entre esses artigos e seus autores, produziu o que Kessler *apud* Weinberg (1974) postulou enquanto um número de artigos científicos que carregam uma relação significativa entre si, de acoplamento, quando tem uma ou mais referências em comum. Para o autor

uma associação entre dois documentos, estabelecidos quando se tem em comum uma alta proporção de palavras-chave, descritores, citações ou outras indicações simples do que lhes dizem respeito com os propósitos desta revisão, as citações serão a única unidade comum (WEINBERG, 1974, p.189, tradução nossa).

Para ele a bibliografia de artigos técnicos é uma maneira pela qual o autor pode indicar o ambiente intelectual dentro do qual ele opera, e se dois trabalhos mostram bibliografias semelhantes há, de fato, uma relação conjectural entre eles. Assim como Weinberg (1974), GRÁCIO (2016, p.85) pontua que "o conjunto de referências utilizadas pelos autores em seu artigo evidenciam o ambiente intelectual no qual trabalham e, se dois artigos apresentam bibliografias similares, há uma relação implícita entre eles".

Na avaliação de WEINBERG (1974, p.191, tradução nossa) sobre os estudos de Kessler, "a literatura da ciência é mais do que apenas uma loja de informação factual, tem uma lógica e dinâmica que podem sugerir interconexões que aparentemente não são óbvios à primeira vista". O autor sugere então que o acoplamento bibliográfico pode ser usado para sinalizar quais artigos devem ser lidos por quem e, ao nosso ver, essa é uma medida fundamental, pois ao destacar artigos referenciais se cria um ciclo virtuoso de citações. Como conseguinte esses artigos e seus respectivos autores terão em sua influência um peso destacado e indissociável para composição do campo.

Hjørland (2013) atribui a esse fenômeno ao que denomina enquanto identidade de autor, pois para ele uma ideia de "identidade de citação de um autor", pode ser

compreendida enquanto perfis individuais dos pesquisadores na seleção de referências para suas publicações. Assim, para ele, entender o acoplamento bibliográfico é, portanto, entender o grau de sobreposição em diferentes autores.

#### 4.5.1 Acoplamento bibliográfico de autores

A medida que duas obras referenciam pelo menos uma mesma publicação, caracteriza-se enquanto fenômeno do acoplamento bibliográfico por este constatar que possuem um *link* entre si. Spinak (1996) presume que, se eles se referirem à mesma fonte, há denotação de um sinal de proximidade temática entre os dois documentos. A intensidade dessa conexão dependerá de quantas referências eles têm em comum. Todavia, Kessler *apud* Weinberg, em seus experimentos constatou que o movimento de análise do acoplamento bibliográfico "é independente de palavras e linguagem e os trabalhos que compartilham uma unidade de acoplamento com um dado artigo podem ser considerados 'referências lógicas'" (WEINBERG, 1974, p.190, tradução nossa). Assim entendemos que existe um ciclo virtuoso de trabalhos e autores, pois a influência na estrutura da ciência praticada pelo acoplamento sofre pouca interferência de análises viciadas pela citação de elementos meramente de senso comum.

No caso da métrica discutida neste tópico, separaremos as autorias dos outros elementos que podem ser acoplados bibliograficamente como forma de construir um cenário complementado pelas redes de acoplamento desses demais elementos - como, por exemplo, o acoplamento de referências e periódicos. Para GRÁCIO (2016, p.86)

a análise de acoplamento bibliográfico de autores (ABA), como uma ampliação do método proposto por Kessler (1963), a fim de mapear as atividades científicas dos autores e contribuir para a obtenção de uma visualização mais real e aprofundada das estruturas intelectuais de um domínio científico e ampliar a compreensão das análises de redes de citação baseadas nos autores. No método de acoplamento bibliográfico de autores (ABA), considera-se que quanto mais referências em comum dois autores têm em suas obras, mais semelhantes são suas investigações. (GRÁCIO, 2016, p.86).

A autora defende que na análise do acoplamento bibliográfico o que está possibilitado é um mapeamento das influências intelectuais, complementado a nosso ver, portanto, o mapeamento da estrutura intelectual do domínio possibilitada pela análise de

citações. Com vistas a produzir pelo menos um esboço dessa estrutura, analisamos os 3.453 autores recuperados na revisão de literatura e, para produzirmos a estrutura passível de análise, perseguimos os autores com no mínimo 4 publicações que, pelo menos, tenham sido citados em 40 outros trabalhos ou seja uma representatividade de dez vezes seu número de publicações.

Destes chegamos a um núcleo de 14 autoridades cuja relação se pratica pela quantidade de mesmas referências entre si, dessa forma temos o seguinte grafo dessa correlação

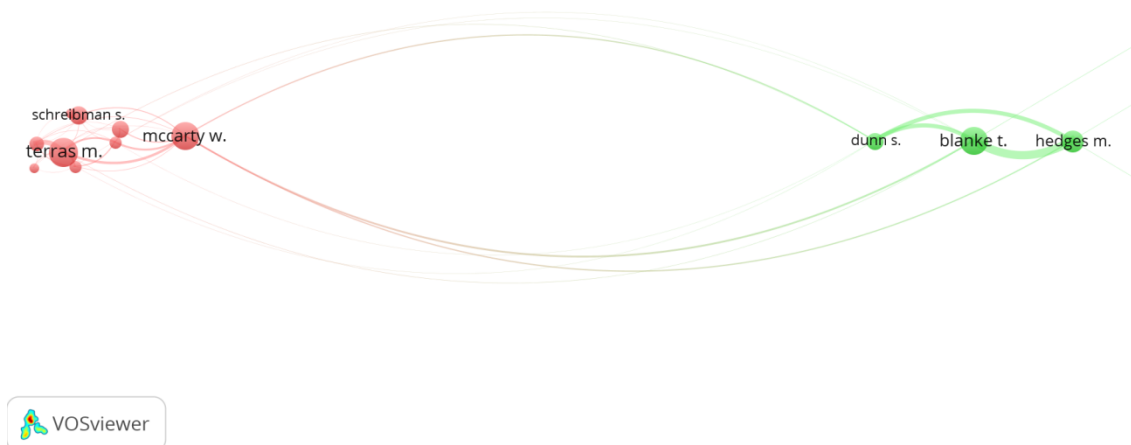


Figura 16: Grafo do acoplamento bibliográfico dos autores.

A rede do acoplamento dos autores sugere uma importante constatação, pois o mesmo se baseia em 3 *clusters*<sup>45</sup> sendo que o primeiro se impõe como central com a destacada concentração de autores. O terceiro *cluster* (não exibido na imagem) apresenta apenas 2 autores que praticam intensa relação de acoplamento entre si e por isso nos furtamos a sua interpretação por julgar que estes pouco contribuiriam para o contexto geral. O segundo *cluster* (de cor verde) se compõe por 3 autores, que mesmo tendo pouca concentração de autorias apresenta intensa proximidade com o agrupamento principal. Todavia este *cluster* representa um agregamento de indivíduos de áreas acadêmicas semelhantes, muito embora por indivíduos tradicionalmente reconhecidos por suas atividades em coautoria (ver 4.3).

Paralelamente a isso o *cluster* principal se destaca como central por ter um número maior de componentes, mais que isso sustenta inter-relações homogêneas e principalmente entre indivíduos de distintas áreas acadêmicas.

<sup>45</sup> O terceiro *cluster* foi suprimido da exibição por prejudicar a mesma.

Tabela 9: Acoplamento bibliográfico por autor.

author	documents	citations	total link strength
blanke t.	13	95	286
hedges m.	10	66	278
conlan o.	10	44	245
hampson c.	9	43	245
terras m.	14	200	180
dunn s.	7	40	153
mccarty w.	13	68	123
warwick c.	6	174	105
vanhoutte e.	5	49	68
nyhan j.	5	56	44
hockey s.	7	61	21
schreibman s.	8	48	16
thelwall m.	4	174	7
nowviskie b.	4	40	3

Fonte: VOSviewer.

O *cluster* mais significativo desta rede, apesar de não possuir uma hegemonia na área acadêmica dos indivíduos, tem as autorias provenientes da área informacional exercendo sua influência já que autores como 'terras m.' e 'warwick c.' serem, por exemplo, os maiores responsáveis por publicar e também aqueles com maior número de citações. Concomitantemente desponta 'thelwall m.' como a terceira autoria mais produtiva e citada, o ineditismo deste ator agrega, à problematização dessas análises, pela participação enfática de um indivíduo advindo das temáticas da altmetria. Imprescinde também destacar que esses três atores possuem números de atividade bibliográfica muito superior, por exemplo, aos primeiros colocados ranqueados por força de *links*.

Apesar do pouco destaque apresentado pelo *cluster 2*, sobressai as primeiras colocações de seus membros. Como os 'nós' da rede correspondem aos indivíduos mais citados em acoplamento, podemos sugerir que esses dois autores são essenciais para literatura das humanidades digitais por exercerem influência direta na estrutura intelectual do campo. Dessa forma sugerimos que este agrupamento representa o qualitativo mais retrospectivo da influência intelectual na literatura das hd's enquanto que a primeira *clusterização* esboça a expansão desta com novos membros e áreas advindo de temáticas também diversificada.

#### 4.5.2 Documentos referenciais na análise de acoplamento bibliográfico

A lógica de que as bibliografias de artigos técnicos são uma maneira pela qual o autor indica o ambiente intelectual em que ele opera, pois se dois artigos mostram bibliografias semelhantes, existe uma relação implícita ainda maior entre eles, essa relação - intelectual - também se aplica aos documentos em si. Ou seja, podemos dizer que não apenas os autores, mas também suas obras expressam uma influência intelectual proveniente das referências citadas em comum. De toda forma

essa sobreposição pode ser parcialmente determinada por diferenças entre os diversos domínios científicos, uma vez que em alguns deles, os autores têm grande liberdade na escolha dos problemas de pesquisa, métodos de investigação e também no que considera literatura relevante. Por outro lado, em outros domínios, os pesquisadores estão bastante restringidos pelas normas e convenções desenvolvidas coletivamente. Desse modo, as identidades de citação apresentam maior variabilidade em alguns domínios do que em outros. (GRÁCIO, 2016, p.86).

Para a autora essas influências de citação em consonância de dois ou mais itens podem revelar alguns cenários peculiares como a utilização de autores clássicos, todavia essa é uma constante apenas para alguns domínios científicos enquanto em outros existe um desprendimento maior na utilização de literatura clássica. Uma outra percepção é que em domínios menos amplos exista uma restrição ou uma menor diversidade de literatura que minora essa utilização. Para nós esse é o caso das humanidades digitais que, por sua incipiência, acaba por concentrar - ou centrar relevância - em publicações expoentes que vão, de fato, representar tamanha relevância para o campo. Assim, nestes casos, temos a condição de inferir que estas publicações são ainda mais representativas do que em campos mais desenvolvidos epistemologicamente.

Hjørland (2013) afirma que uma disciplina acadêmica é tanto um corpo de conhecimento intelectual como uma unidade social tendo seus aspectos intelectuais do conhecimento organizados em conceitos, proposições, modelos, teorias e leis. Por essa razão a manifestação dessa configuração deve se desdobrar de modo especial nas humanidades digitais por ser um domínio incipiente que carece de identidade epistemológica. Assim julgamos que os resultados proeminentes dessas análises, como a

que propomos neste tópico, têm alta capacidade de refletir um cenário real devido à peculiaridade do campo.

Com isto processamos os 1.831 documentos disponíveis para esta análise, exigindo que destes apenas os que possuíssem pelo menos 77 citações compusessem o núcleo de análise. A rede que se forma compõe-se de 10 *clusters* dos quais apenas o primeiro (vermelho) integra-se por mais de um documento e os dois únicos clusters a estabelecer relação são o primeiro (cor vermelha) e o quarto (cor amarela). Dos 10 documentos filtrados, apenas três mantém inter-relação e por isso, formam o *cluster* 1 (vermelho) como podemos verificar abaixo:

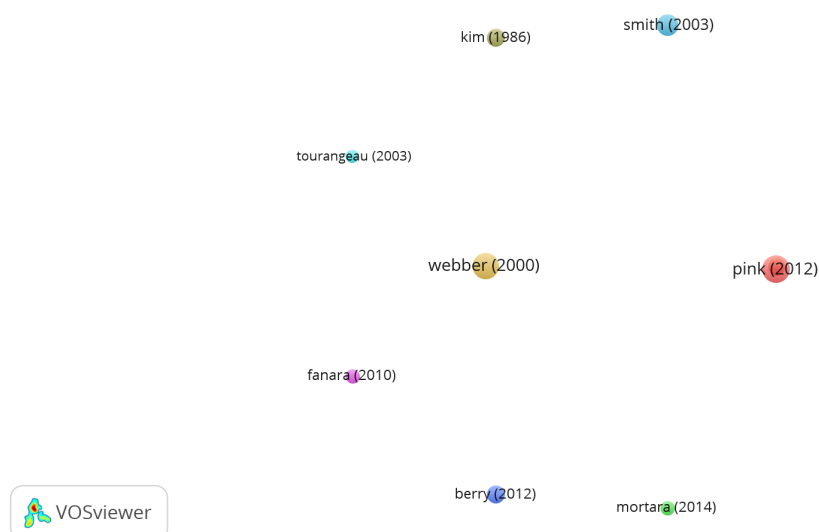


Figura 17: Grafo do acoplamento bibliográfico de artigos.

Após uma tendência de concentração nas análises de métricas anteriores, que consideramos um expressivo indício de convergência teórica do domínio, o grafo de rede desta métrica propõe um movimento oposto, consonante com a ideia de pulverização do campo, atestada neste estudo pelas análises por citações realizadas nos primeiros tópicos. Como este é um retrato passado da publicação no campo, deve refletir o movimento inicial de composição das humanidades digitais, que versa sobre o carregamento de diversas áreas em torno de um rótulo para práticas em comum, que encontra sentido na pulverização dos artigos e na quase nula existência de conexão entre eles.

Nessa medida observamos o *cluster* 1 sendo composto por dois documentos, dentre eles: '*Situating Everyday Life: Practices and Places*' o mais citado e o '*Enabled backchannel: conference Twitter use by digital humanists*'. O primeiro artigo provém da área da teoria da cultura, não fornecendo muitos subsídios para afirmação de proveniência

de uma área claramente delineada. O segundo, todavia, trata de um estudo de usuário em redes sociais e fora publicado no '*Journal of documentation*'. Estes se relacionam apenas com o *cluster* 4 (não passível de visualização na imagem acima devido sua proximidade com o *cluster* 1) com uma obra publicada na '*Scientometrics*' também sobre análise de Twitter.

Tabela 10: Acoplamento bibliográfico por artigo.

document	citations	total link strength
pink (2012)	187	1
webber (2000)	181	0
smith (2003)	141	0
kim (1986)	111	0
berry (2012)	110	0
ross (2011)	103	2
fanara (2010)	89	0
mortara (2014)	88	0
holmberg (2014)	87	1
tourangeau (2003)	77	0

Fonte: VOSviewer.

Os demais *clusters* não apresentam expressividade, à exceção daquele formado por 'webber (2000)', cujo documento também fora publicado pela '*Journal of information science*'. Este documento, apesar de integrar um *cluster* de apenas um documento como os outros nove, se destaca por sua quantidade de citações (a segunda maior), dissonante dos demais. A publicação deste documento num periódico da área de '*information science*' reforça sua proximidade temática com área ao abordar uma temática própria desta. Ao nosso ver o artigo '*conceptions of information literacy: new perspectives and implications*' trata de uma temática da área da informação sob abordagem de novas reflexões com o que vem se conceituar enquanto humanidades digitais.

Apesar de a interpretação da análise dessa métrica se configurar como complexa, pois não fornece subsídios parametrizados em equilíbrio entre o universo de *clusters*, é sensível compreender o papel da área de '*information science*' nas proveniências acadêmicas dos artigos destacados por acoplamento bibliográfico.



#### 4.5.3 Periódicos em acoplamento bibliográfico

Com uma análise menos restrita ao nível dos artigos, o acoplamento bibliográfico de periódicos revela relações mais evidentes sobre o ambiente intelectual nos quais estes vetores da literatura acadêmica trafegam. Aplicadas também no âmbito da classificação de periódicos, como defendem Small; Koenig (1977), consiste na formação de agrupamentos de campo ou especialidades para estabelecimento de alguma medida de similaridade entre eles para, deste ponto, se estabelecer uma comunidade discursiva.

Para Boyak; Klavans (2010) o processo geral para o acoplamento bibliográfico se estrutura na identificação de um conjunto de periódicos relevantes, no cálculo da semelhança entre pares usando contagens bibliográficas de acoplamento e na atribuição de trabalhos de citação a grupos usando os valores de similaridade. Nesta análise, viabilizada pelo software, a similaridade é externada pela proximidade entre os itens que indicam a relevância dos periódicos e suas respectivas áreas científicas para a comunidade discursiva que buscamos identificar. Assim, enquanto vetor da literatura científica produzida, os periódicos têm papel fundamental tanto quanto item de análise e principalmente como indicador da percepção do desenvolvimento de um campo.

Dessa maneira partimos nossa análise sobre o total de 892 fontes<sup>46</sup>, e por estarmos focados na relevância desses elementos, elegemos todo o universo de periódicos que publicaram pelo menos um artigo, mas que tenha sido citado pelo menos 100 vezes. Essa escolha nos garante que todas fontes sejam analisadas, não apenas por sua baixa assiduidade em publicação que descartamos em métricas anteriores, mas sim por seu papel junto aos pares representado pela quantidade de suas citações. Com esse fator de corte sobre o valor que mais se aproxima do grupo dos principais, foram aqueles citados pelo menos 100 vezes que nos entregam um recorte com destacados 15 periódicos.

Desta seleção a rede que se forma consiste-se em 4 *clusters*, dos quais todos mantêm ligações entre si. Esse é um fator que merece ser destacado, pois a acoplagem anterior - por artigos - mostrou-se difusa, sem relação entre os 'nós' do grafo, característica que mencionamos refletir a pluridisciplinaridade das *digital humanities*, sobretudo se considerarmos que esta análise aborda um espaço temporal preteritamente consolidado. Enquanto que na análise desta métrica - de periódicos -, percebe-se que apesar de existir uma dissipação da literatura em humanidades digitais por quase 900 títulos, a consistência

---

<sup>46</sup> Nesta análise outros itens que não apenas Periódicos podem ser considerados como tal para análise desde que sejam passíveis de identificação pelo VOSviewer através de sua citação em referências bibliográficas.

do papel destes na vetorização da influência acadêmica no campo pode ser atestada por seus índices de citações - razão pela qual priorizamos este atributo na constituição deste grupo de análise. Dessa forma temos, a seguir, o grafo de rede do acoplamento bibliográfico de periódicos:

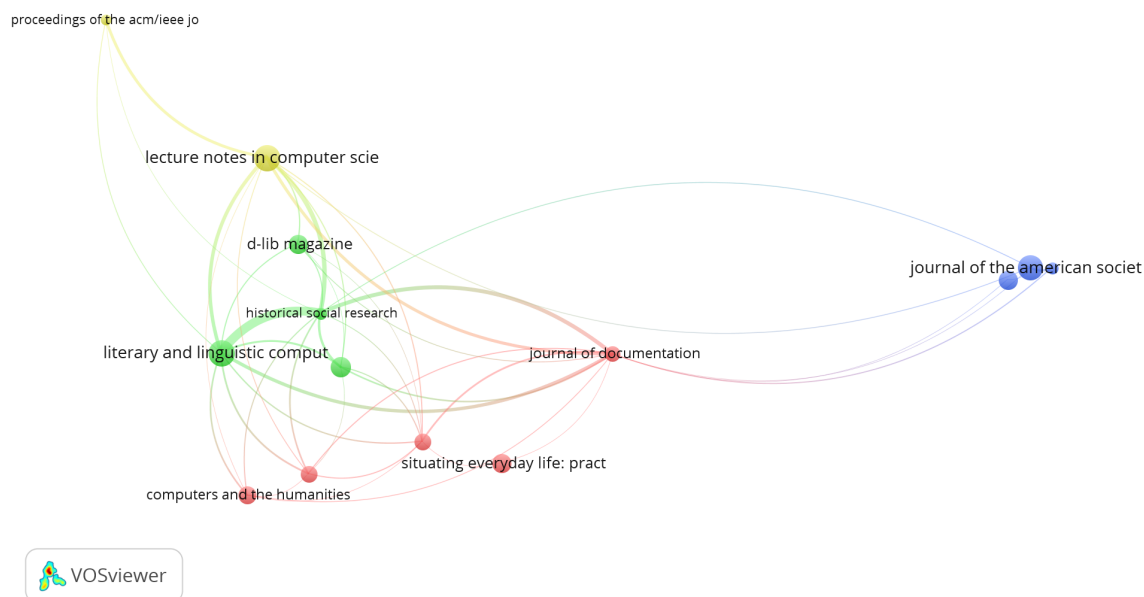


Figura 18 Grafo do acoplamento bibliográfico de periódicos

O principal *cluster* (de número 1, na cor vermelha) reúne o maior número de itens com 4 elementos. Neste é possível identificar um título da área de computação (*computers and the humanities*), também a obra clássica '*understanding digital humanities*'; além destes identifica-se uma publicação bibliográfica que se autotemática como teoria cultural (*situating everyday life: practices and places*). Por fim lista-se duas obras da área de bibliotecas/informação: '*journal of documentation*' e '*library quarterly*'. A constatação da presença desses dois itens no *cluster* 1 nos permite dizer que a área acadêmica de maior relevância - e sem paralelo - é a de bibliotecas/informacional. Essa característica também se confirma no *cluster* 2. Com 4 itens, dos quais 2 também pertencem a mesma área disciplinar dos estudos informacionais: '*journal of library administration*' e '*d-lib magazine*'. O *cluster* 2 apresenta ainda dois títulos dos quais um da área de História e o outro da área de Linguística. O terceiro *cluster* (cor azul) também

corroborar a área dos estudos informacionais como a mais relevante e sem paralelo no acoplamento bibliográfico entre periódicos. De seus três itens, todos pertencem a este campo: '*journal of the american society for information science and technology*', '*journal of information science*' e '*scientometrics*'.

Tabela 11: Acoplamento bibliográfico por periódico.

source	documents	citations
lecture notes in computer science (including subseries lecture notes in artificial intelligence and lecture notes in bioinformatics)	82	283
literary and linguistic computing	44	280
journal of the american society for information science and technology	6	260
journal of library administration	9	208
journal of information science	2	188
situating everyday life: practices and places	1	187
d-lib magazine	8	184
computers and the humanities	20	181
library quarterly	4	168
understanding digital humanities	7	159
journal of documentation	5	150
scientometrics	3	113
ieee transactions on computers	1	111
historical social research	13	106
proceedings of the acm/ieee joint conference on digital libraries	19	102

Fonte: VOSviewer.

Uma outra observação que cabe ressaltar é que os periódicos com maior número de publicação de artigos tendem a equivaler em mesmo patamar de relevância por índice de citações. Com exceção dos títulos '*computers and the humanities*', o terceiro com mais publicações; '*historical social research*', quinto maior em publicações, os itens seguem a lógica de equivalerem sua produtividade com seus níveis de citações. Todavia há um comportamento dissonante nesta afirmação: o caso dos títulos cuja temática contemplam os estudos informacionais (sinalizados na tabela acima). Estes títulos apresentam número de publicações inferior aos primeiros ranqueados e, mesmo assim, veiculam-se como bem

posicionados quando observados por quantidade de citação. Esse fenômeno explica a atribuição de relevância alta à esses títulos pelo algoritmo do *software*.

A natureza dessa métrica contribuiu para permitir uma identificação mais clara das áreas em evidência na literatura das humanidades digitais. O fato dos periódicos servirem de balizadores conceituais contundentes nos auxiliou nessa tarefa, e segundo Mueller (2007), por serem instrumentos de comunicação formal dos resultados da pesquisa original para a comunidade científica servem como arquivo das ideias e reflexões dos cientistas, dos resultados de suas pesquisas e observações sobre os fenômenos da natureza. Diferentemente do caso do acoplamento de referências que serviu para exibir um cenário difuso, a acoplagem de periódicos complementa essa abordagem e concede elementos suficientes para verificar a proeminência da área informacional na vetorização da literatura científica das humanidades digitais.

#### **4.6 Compreendendo a estrutura temático-conceitual do domínio: analisando as palavras-chave**

Para Leung et al (2017) o uso da bibliometria permite que os pesquisadores examinem o corpo da literatura em sua área de estudo para identificar os principais temas, sendo frequentemente combinada com técnicas de mapeamento científico para visualizar a estrutura intelectual de um campo de pesquisa específico.

Com o objetivo desta pesquisa pautando-se pela busca de uma noção sensível daquilo que pode revelar a influência e, principalmente, da presença do corpo acadêmico de outras área do conhecimento na composição das humanidades digitais, não nos apresenta mais adequado finalizar essa análise bibliométrica buscando aquilo que de melhor pode representar o que almejamos: a temática conceitual das obras que nos serviram de estudo. Considerando que caminho escolhido enquanto metodologia tenha sido o tracejamento de um contorno concedido pelo estudo bibliométrico aplicado à esta área, alcançamos neste tópico, como um tipo de conclusão incomparável, a análise sobre as palavras-chave desta revisão de literatura.

Neste contexto que vemos surgir a possibilidade de categorização, que baseada nessa análise das palavras-chave, permitiu estabelecer pelo menos um nível inicial de categorização e nossa proposta é, a partir dele, identificar similitudes que nos permitam identificar a presença/participação de outras áreas neste campo, e de modo especial, a da

Ciência da Informação. Para MEIRELES et al (2016, p.87) a categorização é um processo que busca "associações de tudo que lhe parece novo aos grupos de imagens e conceitos já estabelecidos, sendo o indivíduo capaz de reorganizar as estruturas já concebidas". Todavia a base desta ideia está centrada na busca pela estrutura de conhecimento responsável por revelar características importantes sobre os domínios de pesquisa.

Na pesquisa bibliométrica, as palavras-chave das publicações são consideradas os elementos básicos para representar conceitos de conhecimento e têm sido comumente usados para revelar a estrutura de conhecimento dos domínios de pesquisa (Su & Lee, 2010). [...] Nestes estudos, existem duas diferentes abordagens para o uso de palavras-chave das publicações, dependendo dos objetivos da pesquisa: (1) usar todas as palavras-chave para explorar na estrutura características do conhecimento do domínio no nível macro e (2) usando algumas palavras-chave "importantes" para analisar os detalhes dos principais tópicos de pesquisa de um domínio e sua relação no nível micro. Como muitos estudos empíricos são realizados no nível micro, é necessário estudar o processo de seleção dessas palavras-chave "importantes" para a análise bibliométrica. Em estudos anteriores, os pesquisadores se concentraram principalmente na identificação de tópicos de pesquisa (por exemplo, agrupamento de temas de pesquisa e descoberta da comunidade de rede) e interpretação dos resultados. (CHENG; XIAO, 2016, p.212, tradução nossa).

Para os autores esse tipo de análise se revela abrangente e atende, a nosso ver, a proposta dessa pesquisa uma vez que possibilita dois níveis de observação do cenário no qual o primeiro se pode construir uma interpretação sobre o todo -cenário macro, e o segundo que permite revelar detalhes com a obtenção de palavras importantes, que segundo eles revelam o nível micro. Para este cenário, no qual se orientam pela existência de palavras 'importantes', característica essa que atribuiremos àquelas que demonstrarem centralidade destacada. Para eles "palavras-chave gerais podem ser úteis para mostrar uma visão geral aproximada de uma disciplina científica, mas são menos bem-sucedidos na exibição de temas detalhados de um domínio de pesquisa" (CHENG; XIAO, 2016, p.212 tradução nossa). Assim, ao nosso ver, se reforça a necessidade de se obter através do VOSviewer também as palavras-chave de maior centralidade. Essa é uma das razões pela qual demonstramos interesse em verificar as palavras-chaves atribuídas pelos próprios autores, além daquelas aplicadas pelos editores e periódicos.

Como resumiu Choi et al. (2011), os esforços para identificar palavras-chave importantes das publicações através da bibliometria são baseados na popularidade (principalmente usando frequência de palavras-chave) ou em rede (principalmente usando medidas de centralidade). Palavras-chave de alta frequência são geralmente identificadas como temas importantes de pesquisa para análise bibliométrica. Um número

predefinido ou limite de frequência é geralmente usado para filtrar as palavras-chave. (CHENG; XIAO, 2016, p.213, tradução nossa).

Dessa forma, como propomos na análise que se dedicava aos indivíduos - autores, e àquelas que dedicaram-se aos vetores - periódicos, às quais perseguimos a ideia de estrutura elitista da ciência, neste tópico dedicado à análise das palavras-chave, pretendemos fazê-lo buscando através da centralidade atribuída pela rede a ser construída, a importância que destacará tais termos colocando-os em evidência. Com efeito, para os autores, "se um pesquisador pretende revelar os detalhes dos principais tópicos de pesquisa de um domínio e suas relações no nível micro é necessário selecionar algumas palavras-chave como representação de temas importantes de pesquisa no domínio" (CHENG; XIAO, 2016, p.220 tradução nossa). Assim, julgamos ser um método contundente a proposta de obter o cenário genérico das palavras-chave e partir para um nicho mais objetivo, no qual as palavras centrais possuem maior relevância para o domínio revelando-o e permitindo associá-lo às áreas que compõem as *digital humanities*.

Assim sendo, partiremos para análise das palavras-chave em suas três procedências distintas apresentadas pela base Scopus: totais, na qual o software não se baseia na proveniência e considera todas as palavras-chave encontradas; as do índice, que via de regra são aquelas atribuídas por outros indivíduos que não os autores; e por fim, dos autores, que são aquelas aplicadas pelos próprios elaboradores dos artigos.

#### 4.6.1 Uma visão geral das palavras-chave

Do universo recuperado e em análise nesta pesquisa há um total de mais de 7.000 palavras-chaves classificando os documentos. Se elegermos enquanto fator de corte a reincidência de 50 vezes, chegamos a um núcleo de 11<sup>47</sup> 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. Neste cenário forma-se o grafo abaixo que exhibe os termos gerais mais importantes obtidos neste levantamento.

---

<sup>47</sup> Elegemos 11 termos, pois o mais recorrente será 'digital humanities'.

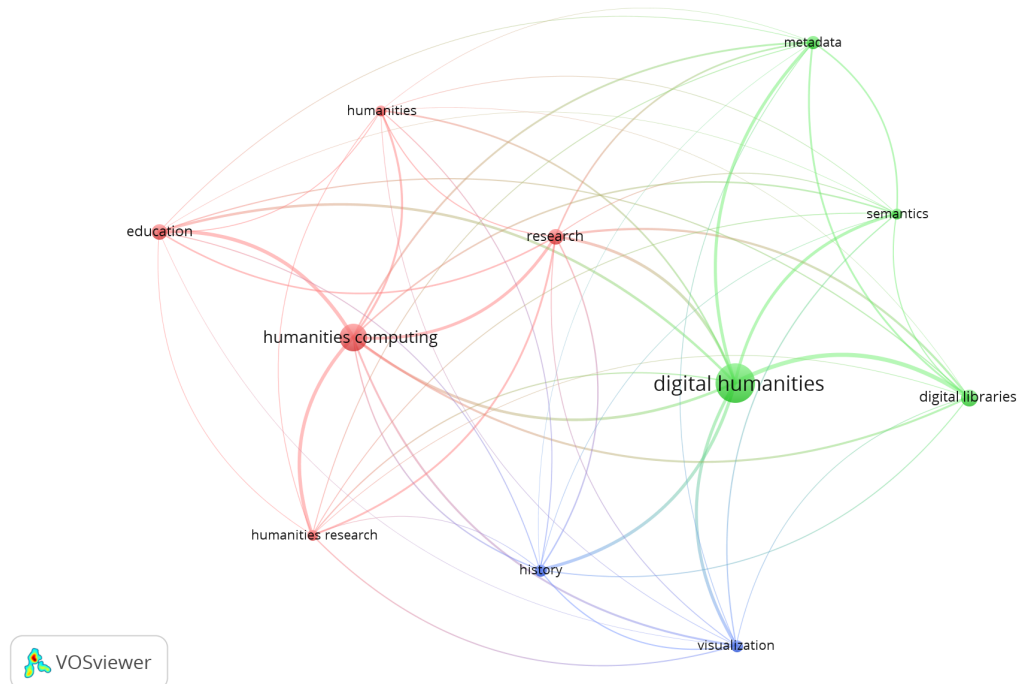


Figura 19: Grafo de ocorrência de palavras-chave totais.

A rede que se forma responde por 3 *clusters*. As áreas acadêmicas são os principais destaques com Educação (*cluster 1*) e História (*cluster 3*), seguidos de termos que podemos considerar orgânicos como '*humanities*' e '*humanities computing*'. O *cluster 2* reúne três itens dos quais '*digital libraries*', '*metadata*' e '*semantics*' sugerem temáticas para os objetos de pesquisa das humanidades digitais. Além destes, o *cluster 3* apresenta uma prática de pesquisa do campo, exibindo o descritor '*visualization*'.

Dos termos principais podemos observar que existe ainda uma ampla abordagem das atividades em humanidades digitais por parte das *humanities computing*, sua antecessora conceitual. Essa constatação pode expressar que ainda há uma relação de proximidade muito alta, refletindo uma dependência conceitual grande, pois se trata do segundo termo mais encontrado. Ao nosso ver, sugere-se que os dois conceitos são pensados conjuntamente, refletindo por parte das humanidades digitais, uma fraqueza epistêmica ou uma necessidade de afirmação acadêmica própria.

Por outro lado vemos enquanto terceiro termo mais recorrente a '*digital libraries*' que vem corroborar as constatações tecidas ao longo das análises das métricas, que demonstraram que a área informacional tinha posição central nos debates uma vez que as bibliotecas eram constantemente abordadas enquanto estudo de caso ou como objeto de

execução das práticas em humanidades digitais, estando quase sempre sendo responsáveis por amparar tais projetos.

Na sequência, em quarta e quinta ordens de recorrência, encontram-se, respectivamente '*research*' e '*humanities research*'. Estas expressões parecem redundar o significado de *digital humanities* que se reverbera na ideia de aplicação de tecnologias para pesquisa no campo das humanidades. Apesar de coerente, essa constatação não apresenta uma novidade sob o ponto de vista da descoberta de temáticas relacionadas.

Na sexta ordem de recorrência, vemos '*semantics*'. Que permite agregar valor à nossa observação devido às práticas de análise de *corpus* textuais ou demais formas de mineração de dados que, sugerimos, possuem relação com esses projetos de análise de grandes volume de textos ou dados por *software* que na maioria das vezes tem objetivo de buscar significações semânticas.

O sétimo termo cita a área de história ('*history*'). Como pudemos ver, essa área acadêmica tem presença marcante no campo, seja como importante aplicadora de metodologias digitais em suas práticas, seja por ser ela própria um objeto desses processos. Em seguida temos '*metadata*' e '*visualization*'. Estes dois termos talvez sejam os mais significantes até aqui sob o ponto de vista temático. Revelam com clareza atividades intrínsecas às humanidades digitais que já podem, sem ponderações, serem consideradas um escopo da área. Todavia as problemáticas respectivas aos metadados não sejam uma exclusividade de uma área específica, temos nos objetos digitais base de uma relação inerente às *digital humanities*.

Os dois últimos termos '*students*' e '*humanities*' mostram-se pouco reveladores. O primeiro pode sugerir o caráter prático da ação dos projetos de humanidades digitais, quase sempre voltados para uma aplicação prática de pesquisa, e neste sentido, tem nos 'pesquisadores' seus principais "consumidores".

Se estendermos o cenário dos termos mais recorrentes para o patamar dos cinquenta mais incidentes, temos a listagem abaixo em ordem de importância atribuída pelo fator '*link strength*':

Tabela 12: Palavras-chave 'totais' mais recorrentes (49 principais).

keyword	total link	
	occurrences	strength
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

Fonte: VOSviewer.

O aumento do escopo para as palavras-chave nos permite observar uma variedade de termos que remetem à temáticas reveladoras sobre o campo das humanidades digitais. Podemos usá-los 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 para práticas 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 nossa 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.

#### 4.6.2 *Palavras-chave dos autores*

Um dos recursos importantes oferecidos pelas bases de dados para análise de domínios de conhecimento é a atribuição de palavras-chaves em suas colocações realizadas pelas diferentes formas de atribuição. Neste tópico nossa análise se baseará nas palavras-chaves dos próprios autores dos artigos. Nos interessa especialmente este recorte pelo mesmo valorizar a abordagem no que os indivíduos tendem a aplicar classificações aos seus textos com grande acuracidade, além de direcionarem suas obras para uma temática e públicos implicitamente declarados.

Nesta análise dispomos de 3.599 palavras-chaves no total. Destas precisamos atribuir enquanto nota de corte para obter as 10 mais importantes, a ordem de recorrência de 22 vezes. A rede se forma com 5 *clusters* dos quais os dois últimos se apresentam enquanto 'nós' únicos: respectivamente o 4º com '*cultural heritage*' e o 5º com '*metadata*'.

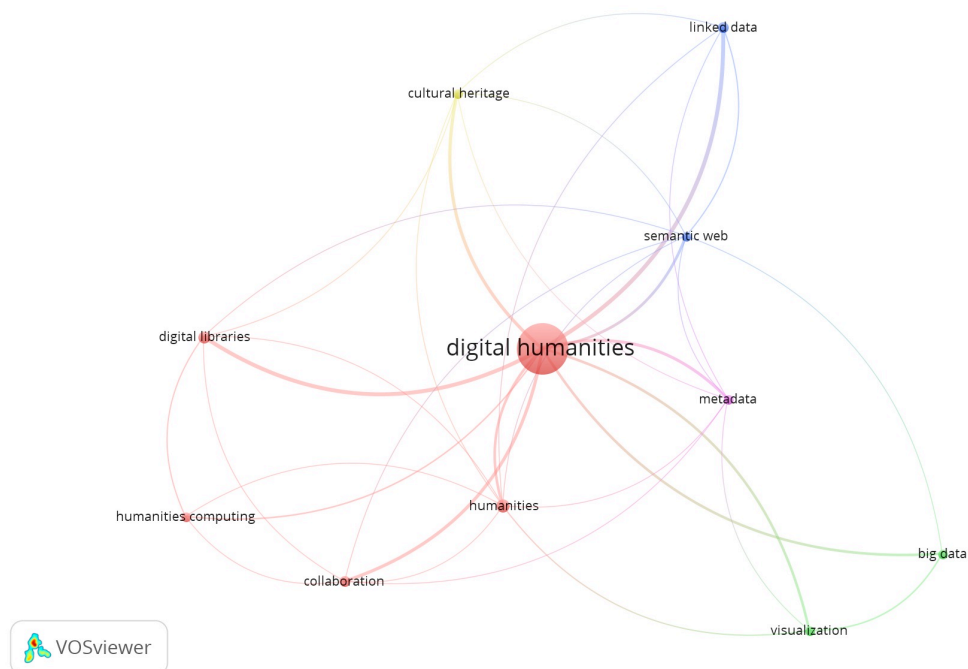


Figura 20: Grafo de ocorrência de palavras-chave de autores.

O *cluster* principal, com 5 'nós', sugere um objeto de estudo (*digital libraries*) conjuntamente com uma prática (*colaboration*) além de uma área '*humanities*'. Também apresenta o conceito de '*humanities computing*'. Excluindo-se o fato de que o descritor bibliotecas digitais seja suficientemente claro, a citação da área das humanidades pouco acrescenta à nossa observação, bem como a presença de '*humanities computing*'. Todavia o descritor 'colaboração/colaborativo' possa ser revelador quanto às práticas desenvolvidas no campo. Os *clusters* 2 e 3 se mostram mais representativos quanto aos objetos de estudos '*linked data*' e '*semantic web*' (*cluster* 3) e quanto às práticas '*visualization*' (*cluster* 2). Esses termos configurariam o nível micro da observação das palavras-chave segundo os autores, tendo a relevância de serem atribuídos por aqueles que, via de regra, têm contundência naquilo que estão classificando por terem eles próprios redigido suas reflexões. Além de que estes indivíduos promovem um enviesamento implícito na atribuição dos termos, impactando diretamente o alvo destes textos.

Para ampliar o escopo de observação e poder analisar o cenário macro destes termos, diminuimos a linha de corte para 9 recorrências, o que nos faz alcançar um universo de 47 termos com maior incidência.

Tabela 13: Palavras-chave 'de autor' mais recorrentes.

keyword	total link	
	occurrences	strength
digital humanities	546	399
metadata	39	55
tei	28	52
ontology	25	49
visualization	40	49
humanities	40	41
digital libraries	31	40
linked data	28	40
archives	27	39
collaboration	24	38
semantic web	22	34
cultural heritage	23	31
digitization	20	29
history	17	26
big data	23	25
text mining	21	25
digital history	19	24
technology	15	24
open access	13	21
crowdsourcing	12	20
education	16	20
humanities computing	23	18
rdf	9	18
culture	9	16
data mining	10	16
digital scholarship	15	16
network analysis	9	16
social network		
analysis	10	16
corpus linguistics	9	15

digital curation	10	15
gis	13	15
research		
infrastructure	11	15
linked open data	9	14
social media	13	14
text analysis	10	14
digital preservation	10	13
evaluation	9	13
research data	9	13
distant reading	9	12
methodology	11	11
art	10	10
data curation	11	10
digital	10	10
e-science	9	10
digital research	17	9

Fonte: VOSviewer.

O cenário macro, construído agora a partir dos termos atribuídos pelos autores, nos permite traçar as três categorias da mesma ordem que aplicamos enquanto categorização, sendo a primeira respectiva às áreas acadêmicas despontadas nesta observação. Dessa forma sugerimos *humanities, cultural heritage, history, digital history, education, humanities computing, digital scholarship e art*. Conforme objetos para as humanidades digitais identificamos *metadata, ontology, digital libraries, linked data, archives, digitization, technology, open access, crowdsourcing, rdf, social network analysis, network analysis, digital curation, linked open data, digital preservation e data curation*. Para as práticas, ou sejam as atividades desenvolvidas nas HDs, que podemos sugerir, compreendem, ao nosso ver *tei (text encoding initiatives), visualization, collaboration, digitization, text mining, crowdsourcing, data mining, network analysis, social network analysis, gis( geografic information systems) e text analysis*.

Dessa forma vemos reincidir neste tópico as áreas acadêmicas identificadas na análise dos termos totais como Humanidades, História e Educação, todavia surgem novas áreas que sugerem ser derivadas da acuracidade com a qual se espera que os autores

classifiquem seus textos. Um exemplo disso é o aparecimento dos termos História Digital e Digital Scholarship. Apesar de não reincidirem aqui áreas relevantes nas demais observações como Ciência da Informação e Ciência da Computação, contudo suas temáticas relacionadas seguem reincidindo como, por exemplo para o caso da Ciência da Informação: digitalização na categoria 'práticas' e em arquivos, bibliotecas digitais, etc na categoria 'objetos'. Tentar interpretar o por quê dos autores terem suprimidos de suas classificações essas áreas, assume-se correr um risco de equívocar-se, todavia não é de se questionar que os autores das humanidades digitais não reconheçam que temas como digitalização, metadados, curadoria digital e preservação digital sejam, prioritariamente, objetos de estudo da área de Ciência da Informação, por exemplo. Ainda se pode destacar para fundamentar esse argumento o aparecimento de termos, categorizados enquanto objetos para as HDs, até então inéditos e que remetem às reflexões da área da informação como *open access e rdf* - problemáticas altamente identificadas com a área de Ciência da Informação. Nosso argumento é de que esses termos estão mais diretamente associados ao campo da informação do que a da computação. Muito embora seja constatada neste recorte a ausência das duas áreas exemplificadas, a presença dessas temáticas, corroboram, ao nosso ver, uma eloquência delas para as *digital humanities*.

#### 4.6.3 Palavras-chave da base de dados

Com a análise dos termos totais nos fornecendo a possibilidade de observação de um cenário amplo, os termos atribuídos pelos autores nos sugerem maior especificidade na classificação dos textos. Por outro lado a utilização de termos atribuídos pelos editores/índices podem fornecer uma abordagem mais abrangente preservando algum nível de acuracidade. Nesse sentido vemos como importante a análise das palavras-chave aplicadas pelos editores considerando o índice temático da base.

Esse recorte de termos se destaca, inicialmente, pela quantidade significativamente superior de palavras-chave encontradas: 4.844. Esse valor é maior em mais de 1000 termos quando comparado ao totalizado pelo de autores: 3.599. Esse dado nos permite ratificar que a abordagem temática dada pelos autores tende a ser mais específica, o que restringe o escopo das palavras-chave.

Partindo para a análise, é necessário aplicar a este total uma nota de corte de reincidência de 36 vezes para alcançarmos aquele que consideramos o nível micro da análise.

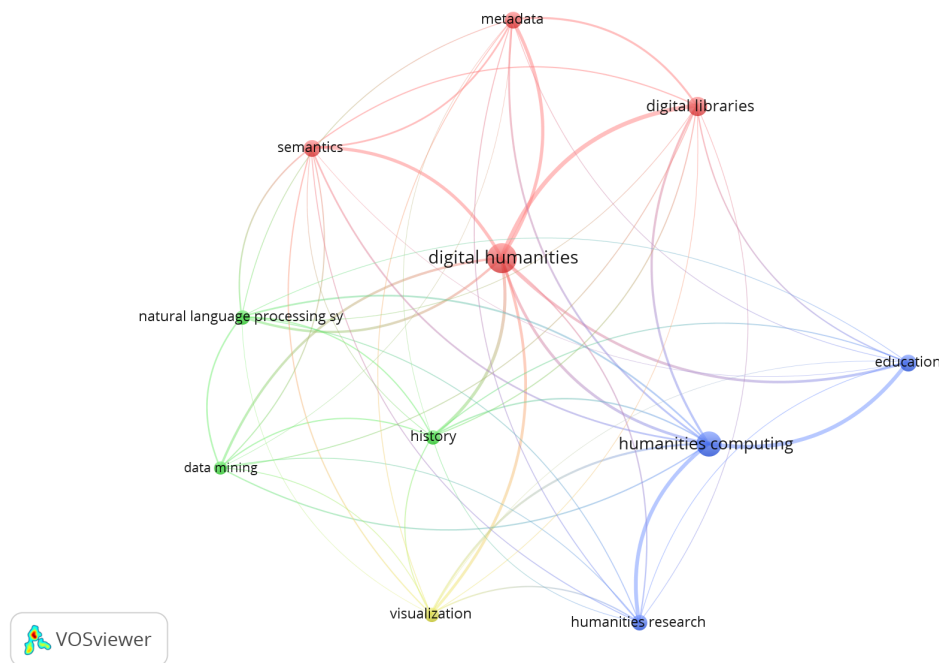


Figura 21: Grafo de ocorrência de palavras-chave do índice.

O grafo de rede destes termos apresenta modificações significativas em relação aos anteriores. Composto por 4 *clusters*, exibe uma equiparação de relevância entre estes. Destes, o segundo e o terceiro possuem 3 itens, e o primeiro 4. Constata-se um certo equilíbrio entre práticas e objetos conforme definidos em nossa categorização, além disso há ainda a emergência de duas áreas que até então não haviam sido identificadas na análise de nível micro: *history* e *education*.

Enquanto práticas de pesquisa nas humanidades digitais observamos '*natural language processing systems*' e '*data mining*' (*cluster 2*), além de '*visualization*' no *cluster 4*. Sobre nossa percepção sobre os objetos de pesquisa, podemos apontar '*semantics*', '*metadata*' e '*digital libraries*' todos no *cluster 1*. Nos importa destacar que os únicos termos a reincidirem em todos níveis micro de recortes de palavras-chaves foram **bibliotecas digitais, metadados, visualização e semântica**. Assim sendo, quanto a estes termos, sugerimos não pairar dúvidas sobre seu nível de representação temático-conceitual da e para as humanidades digitais.



Considerando agora o nível mais amplo, reaplicamos o fator de corte à reincidência de 19 vezes, o que nos entrega um universo de 50 termos. Desta maneira elencamos os 50 termos mais recorrentes na palavras-chave do índice:

Tabela 14: Palavras-chave 'do índice' mais recorrentes.

keyword	total link	
	occurrences	strength
digital humanities	356	780
humanities computing	297	595
digital libraries	106	260
education	93	235
metadata	61	179
semantics	54	160
humanities research	53	142
natural language processing		
systems	35	131
history	57	126
visualization	48	123
semantic web	34	118
information systems	32	112
linked data	29	111
data mining	35	108
social sciences	34	108
information retrieval	34	107
computational linguistics	32	105
cultural heritages	34	101
social networking (online)	34	100
data handling	23	99
ontology	29	88
artificial intelligence	36	87
big data	27	82
learning systems	26	82

world wide web	32	82
engineering education	30	80
behavioral research	24	76
ecosystems	28	73
humanities	34	73
linguistics	25	73
arts computing	25	72
computer science	24	69
web services	20	69
design	26	67
education computing	20	66
human computer interaction	27	65
digital collections	19	63
internet	26	62
knowledge management	20	59
digital storage	20	57
information management	21	57
information science	22	56
information technology	29	55
algorithms	20	54
user interfaces	21	53
virtual reality	28	51
research infrastructure	19	49
digital researches	30	46
hypertext systems	19	42
database systems	20	40

---

Fonte: VOSviewer.

Para categorizarmos a observação destes termos em nível mais abrangente, apontamos enquanto áreas acadêmicas identificadas neste recorte: *education*, *history*, *social sciences*, *humanities*, *linguistic* e *information science*. Julgamos serem objetos para as humanidades digitais os termos: *digital libraries*, *metadata*, *information systems*, *linked data*, *information retrieval*, *social networking (online)*, *ontology*, *big data*, *behavioral research*, *web services*, *human computer interaction*, *digital collections*,

*internet, knowledge management, digital storage, information management, information technology, algorithms, user interfaces, research infrastructure, digital researches, hypertext systems e database systems.* Para as práticas de pesquisa consideramos que são apontadas pelos termos: *natural language processing systems, visualization, data mining, computational linguistics, data handling, artificial intelligence, learning systems, engineering education, virtual reality e hypertext systems.*

Pode-se perceber que algumas áreas reincidentem neste recorte como Educação e História, voltando a figurarem Ciência da Informação, Linguística e Ciências Sociais. O que pode revelar que boa parte dos projetos sejam capitaneados nas áreas de Educação e História ou que estejam direcionados para criação de soluções de problemas destas. As práticas de pesquisa seguem, em sua maioria, reincidente como visualização, processamento de linguagem natural, mineração de dados além dos objetos de pesquisa das humanidades digitais, igualmente, também recorrendo nesta análise. Seguem então sugerindo uma eloquência da área da Ciência da Informação revelada pelos termos bibliotecas digitais, pesquisa digital, coleções digitais, armazenamento digital e ontologias.

A análise das palavras-chave em suas três formas de recorte mostram uma considerável parte da estrutura baseado no que esta pesquisa sugeriu no primeiro capítulo, enquanto uma proposta de esquema prévio à concepção das humanidades digitais. Nesta proposta sugerimos que por atividades ou práticas de pesquisa, compreenda-se toda a ação brutalmente impactada pelo movimento computacional, partindo de seu modelo 'tradicional', essa atividade metodológica confronta-se, nesse primeiro momento, com o objeto de pesquisa que, agora sob uma massificação de objetos digitais, acumulados em grandes repositórios de informação conduziu ao encontro esses dois elementos criando bases para uma possibilidade inovadora na qual a análise não mais se recairia sobre o suporte, mas sim sobre conteúdo.

O que se revela pelos termos não é apenas a estrutura no qual se sustenta o novo campo das *digital humanities*, mas principalmente a possibilidade de identificar, além de áreas eloquentes, quais contribuições estas fornecem. Dessa forma observamos o despontar contundente de áreas como História e Educação, todavia sua presença significa mais que estas áreas participem das humanidades digitais por serem consumidoras de seus desdobramentos e, por isso, servem em maior medida de campo de atuação, do que praticarem uma eloquência baseada na agregação de reflexões e problemáticas a serem

aplicadas às HDs. Este aspecto se revela com a quase nula presença de termos que representem temáticas destas duas áreas no que categorizamos como práticas e objetos.

Diferentemente de áreas como Ciência da Informação e Computação, cuja a primeira, por exemplo, fornece reflexões em temáticas que observamos associarem-se, mantendo um compartilhamento de diversas temáticas em comum. A segunda apresenta-se como parte indissociável das humanidades digitais, suas temáticas mostram-se amparadas no desenvolvimento de recursos computacionais que se traduzem em ferramentas aplicadas à metodologia de pesquisa -expressão mais sensível das *digital humanities*. Essas características indicam que problemáticas destas áreas acadêmicas, sobretudo da primeira, marcam forte presença no que categorizamos como práticas e objetos nas *digital humanities*.

#### 4.7 Instituições produtoras da literatura

Uma métrica não usual para estudos bibliométricos, mas que com o desenvolvimento de versões mais ricas em recursos<sup>48</sup> do VOSviewer passou a estar sujeita a observação, é a métrica de instituições produtoras da literatura no que tange sua citação. Agregar à esse estudo um conhecimento sobre quais departamentos influenciam e destacam-se na produção de conhecimento sobre humanidades digitais é igualmente revelador para compreensão do domínio. Assim nos debruçamos neste tópico sobre a análise da área acadêmica das instituições responsáveis pela autoria dos artigos e a métrica aqui focada é a citação destas, indicado portanto, a partir da citação de seus respectivos artigos.

Nesse sentido conseguimos identificar 2.624 instituições<sup>49</sup>, que quando selecionados por quantidade de publicações, recortamos em 18 aquelas que publicaram, pelo menos, quatro artigos. Ainda assim, restringimos deste número de itens publicados, que pelo menos dois tenham sido citados. Neste grafo se formam 18 *clusters*, dos quais nenhum estabelece vínculo com outro. Esse isolamento não é surpreendente devido a peculiaridade da métrica em análise, ou seja, seria incomum constatar *links* mútuos conectando departamentos. Dessa forma exibimos o grafo estabelecido por esta métrica:

---

<sup>48</sup> Nos referimos a versão 1.6.5 para MacOS

<sup>49</sup> Importa esclarecer que esse número muito provavelmente não representará o cenário real de instituições produtoras, pois como não se trata de uma informação compulsória nas referências bibliográficas, apenas aquelas informadas enquanto metadados serão processadas.

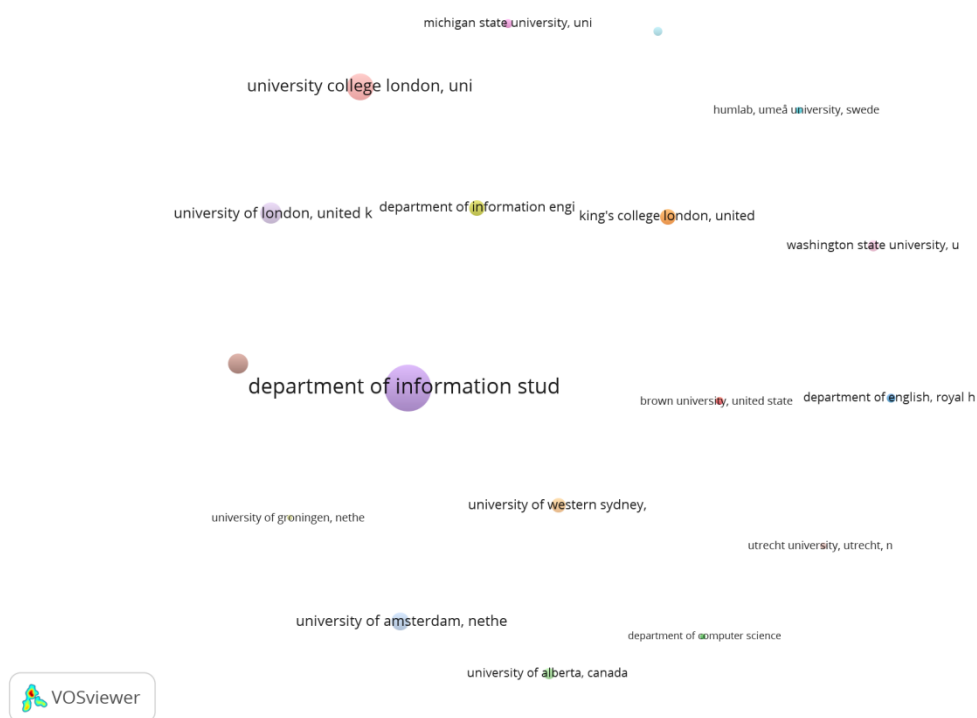


Figura 22: Grafo de instituições produtoras da literatura acadêmica.

A despeito do isolamento dos 'nós' desse grafo, é aviltante o destaque protagonizado pelo departamento de estudos de informação do 'College of London' (*department of information studies, university college london, united kingdom*). Com apenas quatro publicações, esta organização tem com absoluto destaque, o índice de maior citação com 168 notificações. Para se ter dimensão da diferença, o segundo mais citado, identificado pelo VOSviewer como sendo a própria universidade, possui 54 citações mesmo tendo um item publicado a mais (cinco). Na sequência destaca-se ainda que esteja elencado em terceiro lugar a escola de estudos de Biblioteconomia, Arquivologia e Ciência da Informação do mesmo departamento (*school of library, archive and information studies, university college london, london, united kingdom*).

Tabela 15: Instituições ranqueadas por citação.

organization	documents	citations
department of information studies, university college london, united kingdom	4	168
university college london, united kingdom	5	54

---

school of library, archive and information studies,		
university college london, london, united kingdom	4	39
university of london, united kingdom	4	34
knowledge and data engineering group, trinity college dublin, ireland	7	31
university of amsterdam, netherlands	4	25
king's college london, united kingdom	12	20
department of information engineering, university of padua, italy	5	19
university of western sydney, australia	6	18
university of alberta, canada	5	11

---

Fonte: VOSviewer.

O respectivo departamento integra em suas competências os estudos informacionais mais tradicionais como Biblioteconomia e Arquivologia, além de responder aos anseios mais contemporâneos dos estudos informacionais compatíveis com o escopo da Ciência da Informação. Essa característica nos leva a considerar como explicação o fato deste departamento estar sendo considerado na análise de duas formas<sup>50</sup>: a primeira enquanto *department of information studies* e a segunda como *school of library, archive and information studies*.

O outro departamento que pode ser identificado nessa métrica veicula-se em 5º lugar sendo da área de engenharia de dados da universidade irlandesa Trinity College (*knowledge and data engineering group, trinity college dublin, ireland*). O último departamento passível de identificação nesta métrica também pertence à área da computação. O departamento de engenharia da informação da universidade italiana de Pádua (*department of information engineering, university of padua, Italy*) ranqueia-se como o antepenúltimo da lista, publicando cinco documentos e sendo citado quase quatro vezes mais. A presença das engenharias da computação e informática corrobora os dados

---

50 Cremos, portanto, que a segunda forma referida na análise pode ter relação com artigos mais antigos publicados pelo departamento e, a primeira, refere-se a uma nomenclatura mais recente como indica o site do departamento na internet: <https://www.ucl.ac.uk/information-studies/about-department-information-studies>. De toda forma, mesmo constatando se tratar de um mesmo departamento sendo indicado de duas maneiras, não desambiguamos o mesmo na análise por não podermos precisar qual a abrangência de cada um deles.

revelados na análise, todavia surpreende o nível de destaque em questão, visto que a métrica discutida ranqueia-se por citação e a literatura proveniente dessa área de conhecimento não apresentava grande destaque em citações, apesar de corresponderem a importante parcela da literatura levantada nesta revisão.

Por fim, a métrica joga luz sobre a ausência da relevância norte-americana e constata a predominância europeia neste quesito por mais que esse recorte não seja totalmente representativo. Todavia através dela se pode constatar a eloquência considerável de instituições australiana (*university of western sydney, Australia*) e canadenses (*university of alberta, canada*), provocando questionamento sobre a relevância da produção norte-americana nas humanidades digitais.

#### **4.8 Abrangência da literatura**

A literatura acadêmica de um campo é, como defendemos, a expressão mais representativa daquilo que uma área do conhecimento pode significar. Estão nela impressas as digitais necessárias para compreendê-lo uma vez que ela permeia todos os processos de comunicação científica, abrangendo também a influência cognitiva que essas publicações não apenas representam, mas, principalmente, promovem. Todavia as formas de recuperação desse produto cognitivo, comumente compilado, merece ser alvo de considerações por sua incapacidade de abranger uma totalidade utópica. Por isso o método proposto neste estudo se baseou nas técnicas mais tradicionais de recuperação da literatura, aplicados aos meios mais difundidos de seu agrupamento. Neste sentido vimos nas bases de dados o recorte mais adequado para cumprimento da proposta aqui levada a cabo.

Apesar de convictos dessa adequação, não nos furtamos questionar a abrangência desse método de recuperação de produção acadêmica, bem como sua incapacidade de representar a totalidade de um campo, de modo especial no caso das humanidades digitais. Como as bases de dados tendem a abranger periódicos mais qualificados, com reconhecimento manifestado em indicadores de qualidade acadêmica, é esperado que esta seja apenas uma representação acurada de um cenário maior. Essa é uma condição inerente às bases de dados e, por tanto, condizente e conforme aos estudos de levantamento bibliográfico e suas deficiências.

Ao se tratar especificamente das humanidades digitais, consideramos ponderar a incidência dessa deficiência no que tange não apenas às limitações de abrangência das bases de dados consolidadas, mas também fatores outros como as restrições geográficas fruto do impacto da acuracidade imposta pelas bases, além da participação diminuta da produção acadêmica sem origem nos países desenvolvidos sócio-econômicamente. Dessa maneira esta pesquisa tem consciência e sinaliza que as bases de dados tendem a desprestigiar a produção oriunda de países subdesenvolvidos, sobretudo no contexto das humanidades digitais por estas estarem sesivelmente defasadas neste contexto. Ou seja, se assumimos um desprestígio evidente nas bases com publicações pertencentes ao modelo menos acurado de prestígio acadêmico – predominado por países subdesenvolvidos, verificamos que as humanidades digitais distanciam-se ainda mais das práticas da área por razões econômicas e, este viés financista é ainda mais danoso à recuperação da literatura possibilitada pelas bases. Assim, se a literatura acadêmica de países subdesenvolvidos tende a ter presença menor nas bases, no caso da literatura das humanidades digitais isso é ainda mais significativo, pois países não desenvolvidos não acompanham nem a evolução informática base das *digital humanities* muito menos o fôlego de investimentos em pesquisa devido às questões sócio-econômicas.

Considerando assim esse cenário, vemos com relevância estabelecer uma análise geográfica a ser praticada pelo levantamento bibliométrico como forma de corroborar essas duas premissas: a fraca produção de países não desenvolvidos, cujo argumento se decanta nesta análise e a facilidade de desenvolvimento do campo baseado na abundância de recursos capitais. Não obstante, ainda, reafirmando nosso entendimento acerca da predileção das bases de dados mais difundidas no cenário acadêmico pela literatura produzida em países hegemônicos tanto acadêmica quanto economicamente. Nesse sentido verificamos as métricas de citação, coautoria e acoplamento bibliográfico focado na proveniência geográfica de países enquanto autores dos itens.

No que tange a citação de países identificamos itens oriundos de 91 países distintos. Como forma de refinar esse universo, selecionamos apenas aqueles que produziram 10 documentos no mínimo e que receberam, pelo menos, 20 citações – valor escolhido por configurar-se o dobro do número de publicações. Esse filtro alcança o total de 24 países que apresentamos no grafo abaixo:



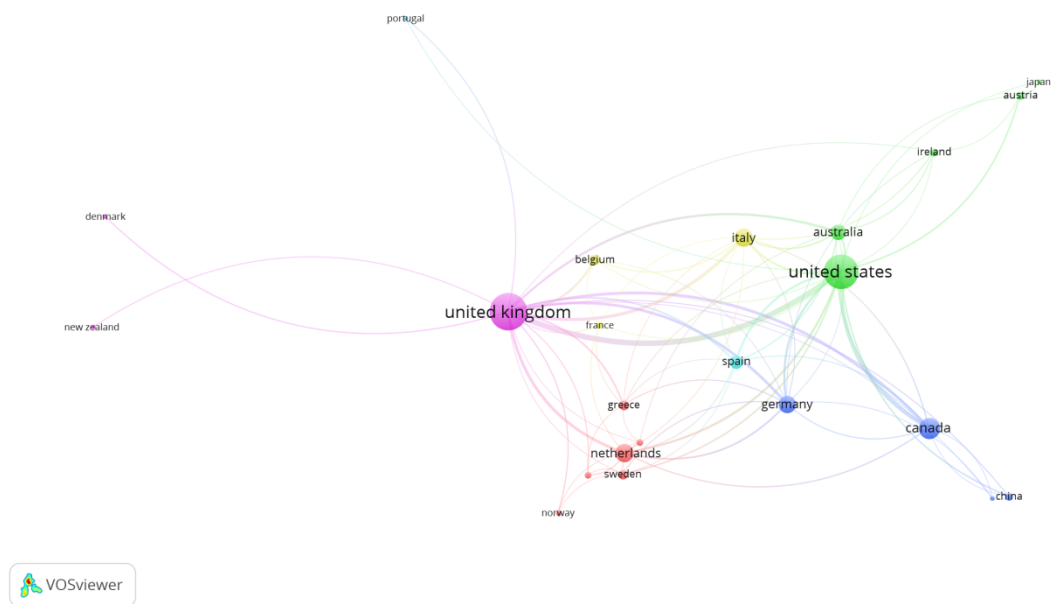


Figura 23: Grafo de citação de países.

O grafo se organiza em 7 *clusters* dos quais, sugerimos serem os 3 primeiros, os principais tendo em vista sua quantidade de publicações e interrelações entre os ‘nós’ no interior do *cluster*. Esses três *clusters* são formados exclusivamente por países desenvolvidos além de China (*cluster* 3) e Rússia (*cluster* 1) que se configuram por sua relevância geopolítico-econômica. Além desses, os *clusters* seguintes 4 (amarelo) e 5 (violeta) - não estão sendo considerados centrais por apresentarem a mesma quantidade de ‘nós’: 3, diferentemente daqueles quais consideramos principais. Destacamos que estes também se constituem integralmente por países desenvolvidos sócio-econômicamente. Os dois últimos agrupamentos *cluster* 6 (verde claro) e 7 (azul claro) tem composição configurada por apenas um único membro, sendo Espanha e Portugal respectivamente. A disposição do grafo corrobora nossa inferência de que as humanidades digitais estão centralizadas nos países detentores de recursos sócio-econômicos. Destacamos ainda, neste recorte, a não incidência de países tidos como subdesenvolvidos, clarificando assim a ideia de que o campo esteja altamente amparado por subvenções econômicas.

No quesito de relacionamento dos ‘nós’ (ou seja entre países) podemos identificar também a forte predominância das nações de hegemonia econômica como Estados Unidos e Reino Unido. Essas duas nações configuram o maior índice de relação mútua, na sequência vemos Estados Unidos e Canadá. O relacionamento entre os outros países

tende a ser equilibrado, não deixando nenhuma margem de destaque. Além disso, chama a atenção que países menos expoentes econômicos como Espanha e Portugal, relacionem-se com EUA e Inglaterra apenas –caso de Portugal. Pode-se perceber ainda que o inglês é a língua não apenas com maior quantidade publicações, mas principalmente a com maior influência pelas citações.

Tabela 16: Países ranqueados por citação<sup>51</sup>.

country	documents	citations	total link strength
united kingdom	296	2225	175
united states	601	3850	149
canada	119	501	60
italy	106	423	45
netherlands	97	542	43
germany	150	444	41
australia	52	564	30
spain	59	180	24
belgium	29	170	15
greece	10	51	13
sweden	33	279	12
china	41	174	8
ireland	35	105	8
austria	23	94	8
france	69	331	7
finland	16	53	7
russian			
federation	12	33	6
norway	15	78	5
switzerland	16	44	4
denmark	17	102	3
portugal	18	21	3

<sup>51</sup> O ranqueamento mencionado refere-se não à quantidade exata de citações, mas ao peso destas que é atribuído pelo índice de ‘Total link Strength’.

new zealand	10	67	2
japan	25	40	2
poland	12	25	0

Fonte: VOSviewer.

Destacamos ainda a relevante produção da Itália, aparecendo entre os maiores publicadores. Fenômeno semelhante ao que se aplica à Austrália ao qual, mesmo não possuindo número de publicações destacado, se projeta com forte evidência quando o ranqueamento se aplica às citações, sendo o quarto país mais citado desse recorte. Outro destaque a ser pontuado refere-se às línguas que não o inglês com alguma relevância nas citações. Caso do alemão, francês, holandês, espanhol e do sueco. Esses idiomas são citados com índices consideravelmente maiores dos que suas quantidades de publicações. Diferentemente desse cenário vemos que o russo, japonês, grego e português são citados com baixa relevância, indicando um desprestígio na influência literária do campo diretamente proporcional aos seus índices de produtividade.

Partindo para interpretação considerando uma ordem de grandeza reversa, podemos identificar que os países subdesenvolvidos enquandram-se todos abaixo da linha de corte de 10 títulos publicados -valor eleito nesta métrica. Apesar de existirem países desenvolvidos acima e abaixo dessa nota de corte, nenhum país não-desenvolvido se encontra acima dela. São exemplos desse caso Colômbia, Chile, Botswana, Argentina, México, Turquia, Cuba, Brasil entre outros. Destes os que mais publicaram foram México e Brasil com 8 publicações cada um, na qual as publicações mexicanas não possuem sequer uma citação, enquanto que as brasileiras, foram citadas 14 vezes.

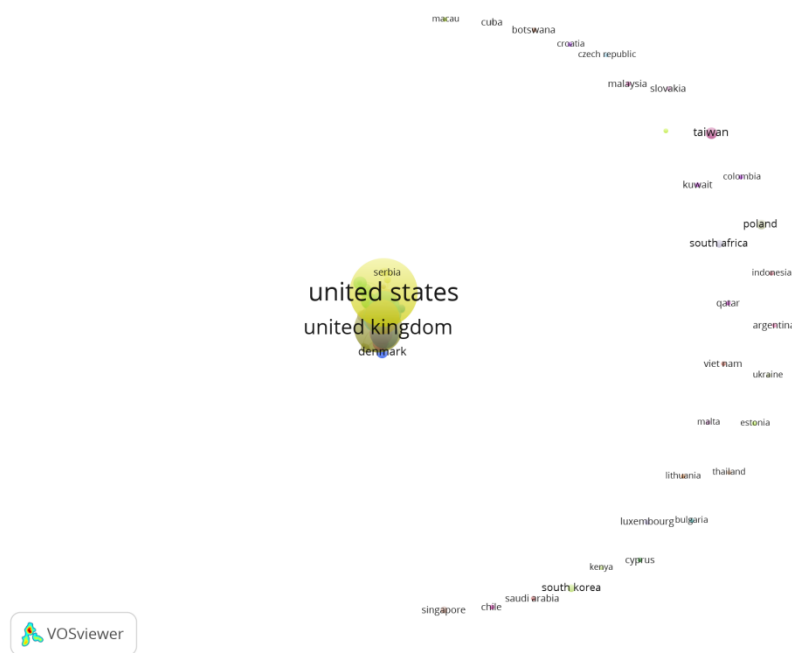


Figura 24: Grafo de citação de todos os países.

O grafo em que constam todos os países citados identificados no levantamento se mostra altamente revelador para uma interpretação sócio-econômica do campo. Nele corrobora-se a inferência de que existe uma centralidade das humanidades digitais nos países aos quais consideramos potências econômicas ou desenvolvidas; mais que isso existe um isolamento destes em relação aos países que produzem menos literatura. Apesar da aviltância dessa apresentação (fig.20) já vem existindo um movimento de aproximação de países subdesenvolvidos com esse núcleo, caso de Brasil e México apenas, por exemplo. Pelo menos uma publicação brasileira se liga a uma publicação norte-americana; já o México se relaciona com publicações espanholas e britânicas. Ao lado de Iran e Índia os dois países latino-americanos são os únicos subdesenvolvidos a se aproximarem, mesmo que irrisoriamente, do núcleo central da produção acadêmica em humanidades digitais.

Nesse sentido, a citação de documentos baseadas em seus países de origem nos serve de argumento para constatar que a literatura das *digital humanities* refletirá uma realidade comum a um grupo seletivo de nacionalidades. Mais que isso, essa realidade considerará o viés da atividade em humanidades digitais num contexto abundante de fomento financeiro. Essa evidência permite apontarmos enquanto mais um fator em sua heurística, que o campo ainda está passível de amadurecimento com as reflexões oriundas de outras realidades que não aquelas nas quais o cenário de existência de fomento é

garantido. De outro modo pode-se especular que há uma menor produção oriunda dessas outras realidades devido a relativa incipiência do campo, todavia o que se discute neste argumento é a discrepância –refletida pelas publicações- entre as publicações de nações consideradas desenvolvidas e as não.

Uma outra métrica nos permite seguir interpretando o campo sob o contexto geopolítico-econômico. Tomando por base a análise de citação dos países, podemos aplicar outras métricas, como forma de validar o argumento sócio-geográfico corroborando-o com a interpretação de outras métricas aplicadas a esse contexto. Dessa maneira aplicaremos também a coautoria de países como forma de extenuar essa análise e incorporá-la à nossa interpretação. Assim, submetemos essa métrica – coautoria- às mesmas exigências aplicadas na anterior: mínimo de 10 títulos publicados e 20 citações recebidas.

A rede que se estabelece nas coautorias permite visualizar um grafo mais consolidado, com uma aproximação maior entre ‘nós’. Composto por cinco *clusters*, fica nítido que existe uma polarização entre os dois maiores produtores da literatura: Estados Unidos e Reino Unido. Podemos identificar ainda um núcleo de países que exerce uma intermediação a essa polarização, estabelecendo ligações com ambos pólos hegemônicos. São exemplos desse núcleo, ao nosso ver, Alemanha, Canadá e Holanda.

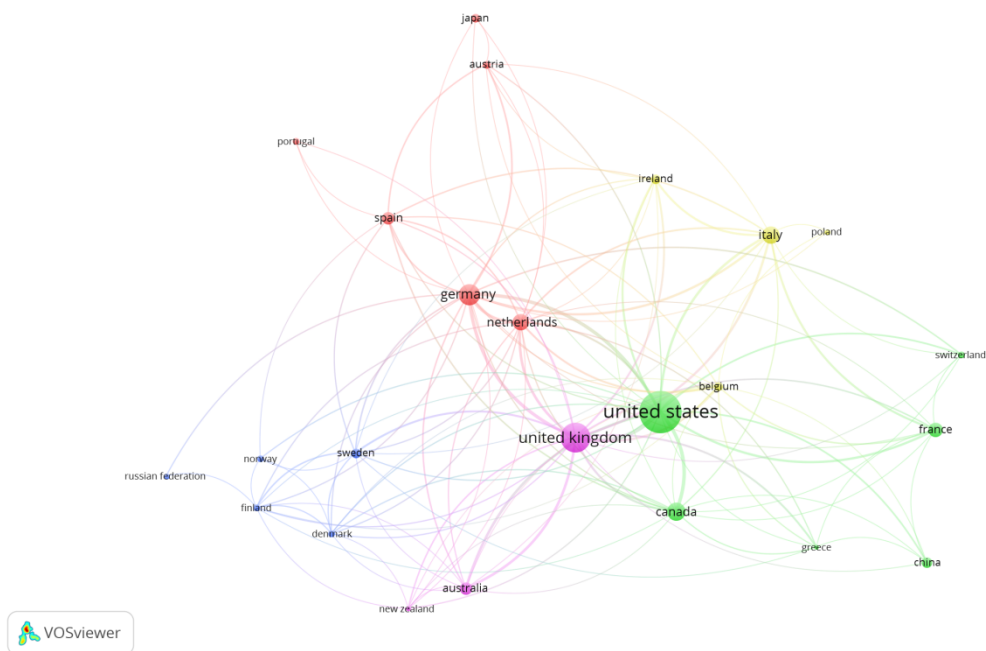


Figura 25: Grafo de coautoria países.

A despeito dessa interpretação, o grafo compõe-se por dois *clusters* principais: número 1 (vermelho) e 2 (verde) compostos por 6 países sendo, respectivamente, Alemanha, Japão, Holanda, Espanha, Áustria e Portugal; e no segundo Canadá, China, França, Grécia, Suíça e Estados Unidos. Os países do *cluster* 1 apresentam um número de produção com destaque, longe portanto do volume publicado pelos pólos hegemônicos, todavia estabelecendo ligações amplas e equilibradas com diversos outros ‘nós’ (países) da rede. Já o *cluster* 2 revela uma participação dos Estados Unidos indicando que este nó hegemônico apresenta uma distribuição maior de ligações entre outros ‘nós’ relevantes. Diferentemente dos Estados Unidos, o Reino Unido mesmo apresentando o segundo maior número de publicações, sugere ligações pulverizadas, perdendo centralidade se observado pelo índice *total link strength* e comparado aos Estados Unidos.

Se o grafo das citações contemplando todos os países indicava com clareza a centralidade da produção e citação de países hegemônicos, projetar o grafo com igual proposta para as coautorias corrobora não apenas o argumento central dessa abordagem, bem como ratifica que o campo das *digital humanities* projeta uma periferia de países não-desenvolvidos. Apesar de não se limitar a eles, todos os países subdesenvolvidos do levantamento se posicionam na metade inferior do ranqueamento, ou seja, não há nenhum país não hegemônico na metade superior desse ranking. A saber:

Tabela 16: Países inversamente ranqueados por coautoria.

<i>country</i>	<i>documents</i>	<i>citations</i>	<i>total link strength</i>
colombia	1	174	0
croatia	1	0	0
czech republic	1	0	0
estonia	1	0	0
georgia	1	1	0
indonesia	1	444	0
kenya	1	13	0
lithuania	1	105	0
thailand	1	0	0
ukraine	1	33	0

---

united arab			
emirates	1	5	0
botswana	2	170	0
cyprus	2	1	0
turkey	2	3	0
cuba	3	2	0
kuwait	3	0	0
qatar	3	3	0
saudi arabia	3	2	0
slovakia	4	542	0
romania	6	2	0
bulgaria	1	14	1
iran	1	51	1
malta	1	0	1
malaysia	2	40	1
serbia	2	4	1
slovenia	5	0	1
india	13	3	1
mongolia	2	0	2
viet nam	2	11	2
argentina	4	8	2
hong kong	7	53	2
brazil	8	14	2
south africa	8	67	2
chile	2	501	3
latvia	3	69	3
singapore	6	3	3
poland	12	14	3
taiwan	18	2	3
macau	1	423	4
portugal	18	3	5
south korea	8	78	6
israel	10	1	6

---

russian federation	12	22	6
mexico	8	5	7
hungary	9	331	7
norway	15	3	7
japan	25	0	7

Fonte: VOSviewer.

A tabela acima lista a metade menos representativa dos países, ilustrando, portanto, que de todos estes, sequer um ocupa a metade mais representativa desse levantamento. Apesar da fraca produção desses países, ratificamos aqui os dois argumentos levantados, o primeiro no tocante à contribuição das práticas desses países para as reflexões prático-teóricas para um campo incipiente e que a produção acadêmica destes além de ser diminuta é irrelevante.

Ainda assim podemos aplicar a análise de acoplamento bibliográfico com a proposta de extenuar a interpretação do papel fornecido pela participação do fator geopolítico-econômico no campo das humanidades digitais. Nesse sentido temos como alternativa buscar um cenário fornecido pelo acoplamento bibliográfico desses países. Como essa métrica utiliza a análise de citações para estabelecer uma relação de similaridade entre documentos, mais que perceber tal similaridade, nos é possibilitado pelo acoplamento bibliográfico, indicar quais elementos possuem hegemonia teórica proveniente da citação de outros dois, haja vista que o que focamos é a indicação de importância na similaridade entre dois a partir de suas citações a um único documento.

Dessa forma procedemos os mesmos dados anteriores aos mesmos filtros de exigência. Neste caso o grafo dessa rede se compõe por cinco *clusters*. Diferentemente dos grafos alcançados nas métricas passadas, este indica que uma consolidação em curso –todavia em nível perceptivo de concretude- pode ser identificado uma vez que existe um *cluster* principal enquanto que, anteriormente, essa distinção não estava passível de ser afirmada. Como o acoplamento bibliográfico ocorre quando dois trabalhos referenciam um terceiro trabalho comum em suas bibliografias, temos uma indicação de que existe a probabilidade de que os dois trabalhos reconhecem e se influenciem da relevância daquele citado. Por essa razão consideramos que o *cluster* 1 reúne tamanha relevância nesse sentido, sobretudo se considerarmos a presença dos Estados Unidos que é o principal produtor da literatura.



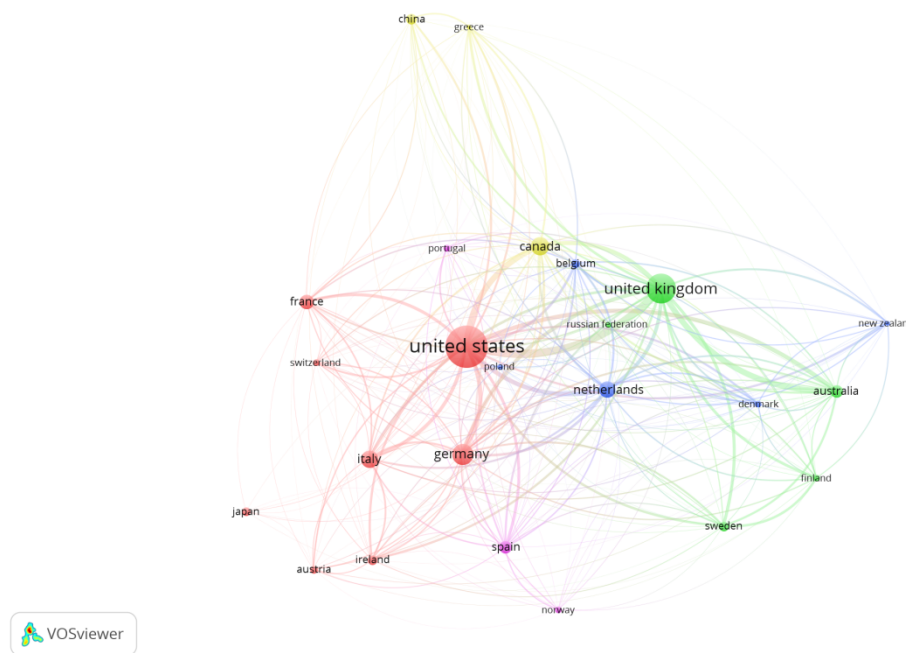


Figura 26: Grafo acoplamento bibliográfico de países.

Percebemos ainda que existe uma forte consolidação entre os dois maiores produtores (Estados Unidos e Reino Unido), contudo passa a existir uma relação considerável de ligação entre estes como indica a espessura do arco conectando esses dois ‘nós’. Apesar dessa forte ligação, a hegemonia do Reino Unido é mantida com sua presença em outro *cluster* (número 2, verde), mantendo também, com isso, sua relevância enquanto citado pelos demais, já que dois documentos são bibliograficamente acoplados se ambos citarem um ou mais documentos em comum. A força de acoplamento de dois documentos é fornecida quanto mais citações para outros documentos eles compartilham em comum, nesse sentido, a relação entre os dois maiores produtores nos leva a inferir que estes são os responsáveis por fornecer tamanha influência teórica suficientemente capaz de configurar o campo conforme suas próprias práticas.

Essa constatação nos permite conjecturar a pouca disseminação das *digital humanities* fora do eixo geopolítico-econômico global, demonstrando ao nosso ver, que as humanidades digitais são um produto da capacidade de financiamento econômico à disposição das áreas acadêmicas humanísticas. Estas, por sua vez, em interlocução com as áreas informáticas, enquanto respostas às duas demandas: de produção de conhecimento baseados em grandes massas de dados informacionais e capacidade criativa de ferramentas que atendam à este cenário. Dessa maneira consideramos o ponto central desta discrepância entre as realidades nesta métrica discutidas.

Por outro lado, este diagnóstico não pode desconsiderar a influência daquela que se apresenta como uma tênue imposição de barreiras pelos meios consolidados de difusão da produção acadêmica, como se apontou anteriormente. Considerando como exemplo da discussão trazida nesta métrica e tomando por base o caso brasileiro, vemos que a literatura acadêmica de países não hegemônicos tendem a ser preteridas por tais repositórios. Pimenta et al (2018), defendem que existe um desprestígio à qualidade da produção brasileira. Segundo os autores esta situação é subproduto de questões semânticas que impactam a recuperação da informação e dos serviços privados de repositórios informacionais que estabelecem padrões de qualidade como filtros de acuracidade. Para eles estas questões impactam diretamente os índices de produtividade no caso brasileiro, que podemos usar de ensaio para outras realidades similares – sobretudo àquelas apontadas na interpretação desta métrica. Os autores trazem relevante contribuição para tomarmos enquanto caso para o discutido, na qual sua experiência de recuperação de uma expressão na base Google Scholar foi capaz de fornecer um número altamente superior de itens. Assim ao utilizar a expressão ‘humanidades digitais’ em língua portuguesa observaram que, segundo o experimento, resultaram pouco mais de 600 itens recuperados em língua portuguesa disponíveis na internet, enquanto que a mesma pesquisa nas bases Web Of Science e Scopus foram reportados 12 itens.

O caso é altamente ilustrativo para as limitações do estudo via metodologia de levantamento de literatura, todavia importa compreender as propostas a que cada repositório se projeta. Apresentando estratégias de indexação e recuperação peculiares, o experimento implementado por Pimenta et al (2018) revela tamanha barreira imposta pelas bases de dados tradicionais consolidadas sob a égide da acuracidade acadêmica - produto de seus padrões de qualidade. Apesar de seu escopo amplamente mais competente, revelando um viés meramente geopolítico, é importante prestigiar não apenas o escopo amplo, mas igualmente, as formas de padronização desses dados como forma de garantir seu processamento com vistas a produção de conhecimento sobre os mesmos.

Dessa maneira esse estudo reconhece que o campo das humanidades digitais tem obstáculos outros além daqueles tidos como geopolítico-financeiros. Tais coibições se furtam a serem analisadas numa abrangência maior devido às limitações do método – cuja limitação é comum a todas as áreas do conhecimento, entretanto consideravelmente mais danosa àquela que este estudo se dedicou e a proposta de sua abordagem.

## CONSIDERAÇÕES FINAIS

Com a experimentação de níveis cada vez mais fundamentados em tecnologias, a área da pesquisa vem estabelecendo a percepção de que este movimento se projeta como um imperativo irreversível. Com grande parte das distintas áreas acadêmicas impactadas pelos vários níveis desse fenômeno das tecnologias da informação e comunicação, um novo estágio se contornou. Como este movimento tratou de um processo abrangente, que influenciou todas as esferas sociais vemos, atualmente, essas transformações serem não apenas percebidas e praticadas, mas interpretadas. Diante da massiva penetração tecnológica que transformou completamente o cenário de convivência científica, o debate sobre como esse processo vem reformatando as práticas metodológicas amparadas nas tecnologias digitais ganha reconhecimento de área acadêmica. Essa lógica daria corpo ao que se cunhou enquanto humanidades digitais.

Dessa forma podemos afirmar que as *digital humanities* representam um patamar de maturação das reflexões a respeito do imperativo do digital considerado enquanto contexto aos debates do que, durante longo tempo, se balizou pela ideia de *humanities computing*. Até então, a penetração de um estatuto das tecnologias da informação e comunicação ditando a partida para um processo de consolidação do uso de recursos dessa natureza, basicamente pela possibilidade de bases de dados voltadas à recuperação da informação, congregava indivíduos de matizes acadêmicas diversas ao redor de uma mesma problemática. Uma reflexão oriunda destes que experimentaram a metodologia tradicional elevada a uma nova dimensão, na qual a tecnologia é a maior responsável pela execução da função árdua nas metodologias de pesquisa, pontua a distinção entre *humanities computing* e *digital humanities*. Assim o intercâmbio de "bens comuns" de natureza metodológica aproximou campos de pesquisa em torno da descoberta de novos métodos informáticos -motivando uma conexão entre campos disciplinares produzindo trocas mútuas - configurando a ideia de *humanities computing*. Ao passo que nas humanidades digitais a proposta é partir dessa ideia, mas, principalmente, fazendo evoluir o trabalho colaborativo em um núcleo –apenas agora comum- de um campo coletivo e plural, abandonando uma percepção isolada do fenômeno como nas *humanities computing*.

Não obstante ao que esta pesquisa considera uma maturação do que era difundido enquanto escopo das *humanities computing*, temos no estatuto do digital a condição *sine qua non* para o crescimento dos debates que conceberam às humanidades digitais um *status* autônomo. A esse estatuto consideramos o impacto causado no

processo de transformação capaz de conferir diversas novas formas de transmissão, armazenamento e acesso a esses registros. Dessa maneira vimos não apenas, mas principalmente na ausência de pressupostos físicos, que esta seria a entidade na qual se encontraria as bases para constituição de uma nova área -das *digital humanities*- a partir do avanço das reflexões propostas pela *humanities computing*.

Por conseguinte as características do suporte digital foram responsáveis por pavimentar o caminho a uma nova estrutura de trabalho no âmbito das ciências, reprogramando totalmente os métodos e instrumentos de pesquisa sobretudo nas disciplinas não íntimas, até então, ao fenômeno das TICs. Assim esta pesquisa classificou enquanto uma ontologia do digital o que se propôs ser uma reflexão a respeito do sentido abrangente deste atributo, como aquilo que torna possível suas múltiplas existências e se apropria dessa concepção para pactuar através dela a transição para o que culminou na concepção de *digital humanities*.

Com efeito o novo cenário posto conclame por reflexões capazes de construir uma amálgama teórica responsável por lastrear práticas meramente tecnicistas, esse mote nos coloca diante do desafio, escolhido por esta pesquisa, de identificar algum nível de composição deste novo campo de saberes outrora rasamente categorizado como pluridisciplinar. Neste espectro esta pesquisa se embasou no pressuposto que a maneira mais contundente de analisar um campo com características fortemente incipientes seria partir daquilo que lhe pertence e lhe representa com absoluta propriedade. Assim uma das grandes noções que esta pesquisa abarcou relacionou-se diretamente com a análise da literatura das humanidades digitais, cuja capacidade de exploração de significação deste produto do campo se torna inconteste.

A representação extraída da literatura desse campo baseou-se, inicialmente, num levantamento da mesma seguindo as metodologias das revisões de literatura. Contudo pretendeu-se aprimorar esse método, conduzindo-o no sentido das revisões sistemáticas. Apesar de não realizar uma revisão dessa natureza, características dessa metodologia foram aplicadas para garantir maior rigidez ao método em questão, explorar criteriosamente termos representativos, minorar a incidência de vieses, além de viabilizar sua reprodução. Esta, assim, se configura mais uma afirmativa desse levantamento, já que um campo incipiente teria essa peculiaridade refletida compulsoriamente em sua literatura. Dessa forma o mapeamento da literatura se norteou por essa metodologia, ciente das complexidades do caso das humanidades digitais, razão pela qual não a propusemos para este caso.

A metodologia aplicada através da literatura das humanidades digitais se balizou nas análises de ordem bibliométrica: estratégia que esta pesquisa julgou adequada enquanto técnica quantitativa para medição de índices de produção e disseminação do conhecimento. Essa estratégia nos possibilitou o estudo e avaliação das atividades de produção e comunicação científica, na qual o levantamento qualitativo dos dados permitiu, senão uma imagem fidedigna do todo, uma possibilidade de identificar nuances que possibilitaram conhecer e investigar características desta área em formação. Assim, através do método bibliométrico, a pesquisa viabilizou buscar as relações entre indivíduos, mapear e identificar os atores, seus lugares de fala e a que assuntos se dedicam no universo analisado.

A relevância deste estudo bibliométrico contribuiu para fornecer uma visão complementar para aquela até então supérflua, tida como meramente pluridisciplinar em seu corolário. Muito embora essa afirmativa se apresente condizente, esta pesquisa buscou senão indicar, pelo menos revelar que o potencial dos estudos bibliométricos baseados em citação seriam capazes de nos fornecer uma imagem menos opaca do domínio das humanidades digitais. No emaranhado de indivíduos que compõem esse domínio, a citação foi tomada como indicador objetivo da comunicação científica, evidenciando as relações extremamente sensíveis entre documentos e seus autores, nos conferindo assim conhecimento menos raso daquele campo antes apenas pluridisciplinar. A narrativa construída nos conduziu à ideia de domínio que representa com propriedade o que a pesquisa se esforçou em apontar: um grupo multidisciplinar se movendo no sentido de pensar pluridisciplinarmente as metodologias digitais de pesquisa.

Para esse grupo multidisciplinar analisado a partir da metodologia bibliométrica, aplicamos, preliminarmente, a estratégia de compreensão e foco de análise baseados em três vetores fundamentais. De fato, para esta pesquisa, a qualidade mais apropriada ao seu desenvolvimento se explicaria e se justificaria nos estudos bibliométricos pela possibilidade de criar pelos menos três eixos de estudo: autores, vetores e representatividade das palavras-chave. Portanto através dos métodos bibliométricos nos foi possível traçar uma estrutura dorsal de identificação do domínio a partir do fornecimento de significados proveniente das autorias, da indicação acadêmica presente nos vetores da literatura, como os periódicos, além do descobrimento da representação temática provida pelas palavras-chave encontradas.

As características pontuadas até então, juntamente com a proposta passível de aplicação, nos garantiram que as humanidades digitais poderiam ser estudadas enquanto

um domínio, pois todo o esforço de ordem bibliométrica conseqüentemente imputa uma análise de domínio ao captar subsídios para mensuração do campo através de sinais oriundos de mapas bibliométricos. A esse processo consideramos a construção de uma análise de domínio de ordem descritiva por ter procedido não a sua definição, mas, sim, a sua descrição. Diante desse processo descritivo entendemos ter visitado imperativos conceituais que viabilizam uma análise de domínio, um; seguindo métodos das Ciências Sociais, dois; se embasando na observação dos atores sociais, e três; por focar num agrupamento alheio às posturas subjetivas como a procedência dos periódicos e palavras-chave, por exemplo. Assim, mesmo se tratando de um campo incipiente epistemologicamente, esta pesquisa conseguiu extrair subsídios essenciais à aplicação de um método possível para construção de conhecimento sobre este complexo domínio das humanidades digitais.

Um dos principais desafios impostos ao estudo das humanidades digitais certamente se apresentou em torno de sua definição conceitual. Como em todo campo de conhecimento científico, a evolução epistêmica do mesmo não deixa de ser impeditiva. Diante da curta existência da criação formal do campo e o impacto que esse *status quo* impõe ao mesmo no que tange sua epistemologia, inferimos que produzir conhecimento sobre esse domínio o enriqueceria conceitualmente. Nesse aspecto esta pesquisa esforçou-se no sentido de alcançar uma definição conceitual próxima das concordâncias já estabelecidas em torno do campo, ao mesmo tempo em que pudesse nortear o estudo bibliométrico e, todavia também ser complementada por ele.

Nesse sentido, como abordado, o argumento desta pesquisa para o conceito das humanidades digitais amparou-se nas reflexões passadas sobre *humanizing computing*. De toda forma o fulcro das reflexões deste estudo centrou-se na tênue distinção entre os dois conceitos. Apesar de não configurarem o sinônimo da abordagem de um fenômeno que, em grande escala se trata de um mesmo, mas que, todavia, se complementa à *humanities computing* a partir da entrada em cena do instituto do digital. Compartilham, assim, um núcleo comum no qual vislumbramos, a partir de então, emergir um gênero de discurso caracterizado por um repertório de temas que oferecem explicações regulatórias e reflete os pressupostos subjacentes sobre seus objetivos de prática. Todavia se configurando no estudo da transformação da escrita, do pensamento e do ensinar nos campos humanísticos promovidos pelas ferramentas informáticas no contexto do digital.

A nosso ver, estamos diante de um discurso que categorizamos em dois momentos necessários para compreensão de um conceito não previsível acerca do campo. Um

primeiro momento prescrevemos, enquanto condição *sine qua non* para a construção do campo das *digital humanities* o tripé das atividades de pesquisa, os objetos destas e as técnicas à estes aplicados. No que seria abarcado pelas atividades de pesquisa, compreenda-se toda a ação impactada pelo movimento computacional, partindo de seu modelo 'tradicional'. Essa atividade metodológica confronta-se, nesse primeiro momento, com o objeto de pesquisa, agora sob uma massificação dos objetos digitais, acumulados em grandes repositórios de informação. O processo que conduziu ao encontro esses dois elementos criou bases para uma possibilidade inovadora na qual a análise não mais se recairia sobre o suporte, mas sim sobre conteúdo, ou seja, seu atributo informacional. Assim abre-se margem ao caráter heurístico das análises do suporte digital por novas ferramentas informáticas. Ao nosso ver, esse é um momento prévio que baliza e sustenta a compreensão das humanidades digitais.

Num segundo momento estamos diante de uma configuração responsável por representar a atuação conceitual do campo, ao qual estabelecemos 4 pontos a serem considerados: a predominância do digital, as práticas desenvolvidas no campo, a agregação de agentes interdisciplinares e a designação da transdisciplina. Vimos, portanto, na predominância do digital, a possibilidade e condição essencial para as múltiplas experimentações metodológicas características primeiras e indelévels do novo campo. No mesmo sentido, as práticas das humanidades digitais fomentaram interesses por diversas práticas, instrumentos, atividades ou ferramentas transversais, provocando no mesmo sentido a agregação de agente interdisciplinares, responsável por novas possibilidades de análise da informação via meio computacional. Por fim a construção de uma narrativa discursiva e auto-reflexiva responsabiliza-se pela designação de uma transdisciplina. Com efeito, esses dois momentos alcançaram aquilo que balizou este estudo enquanto compreensão conceitual das humanidades digitais a designação de uma transdisciplina que congrega em torno de si métodos compartilhados, objetos comuns entre agentes distintos e que vem se identificando como fabricantes de um mesmo produto, pertencente a um escopo comum.

Essa narrativa nos indica a necessidade de buscar os responsáveis por construí-la. Nesse aspecto observamos mais um aditivo de compatibilidade aos estudos bibliométricos. Como uma das estratégias metodológicas para descrição do domínio se volta para compreensão dos seus autores, temos na análise das métricas respectivas a estes mais um grau de importância atribuído. Por essa razão, acompanhamos a centralidade dos indivíduos dessa natureza na relevância da comunicação científica para o

desenvolvimento do conhecimento e principalmente para consolidação de um campo acadêmico. Sendo este um processo que se constitui no conjunto de relacionamentos entre os indivíduos unidos pelos significados que constroem e compartilham entre si, o processo da comunicação na ciência, que se caracteriza enquanto processo vital que desenvolve o conhecimento a partir de elos construídos entre pares, imputando uma impressão de influências que se tornarão marcas daquele determinado campo.

Não obstante, as comunidades discursivas produzidas pela comunicação científica são distintos grupos sociais sincronizados em pensamento, linguagem e conhecimento, constituintes da sociedade moderna. Não são entidades autônomas, mas construções sociais compreendidas por indivíduos. Decerto, ao nosso ver, antes da percepção de uma comunidade científica, há a constatação de um estatuto social que congrega os indivíduos em torno de uma mesma causa. Essa relação involuntária, todavia altamente social, é responsável por conceder às ciências, aos seus processos de comunicação e desenvolvimento o estabelecimento de uma estrutura mais concentrada ao qual este estudo se referiu como 'estrutura elitista da ciência'. Isso porque a política científica tende a ser seletiva e elitista, aonde existe uma alta concentração de produção científica relevante em uma minoria de pesquisadores. Essa estrutura também pôde ser confirmada no domínio das humanidades digitais.

Para este estudo essa percepção foi ainda mais valiosa visto que, devido às peculiaridades das humanidades digitais, a persecução deste ativo acadêmico se resvalaria também enquanto ativo social dando margem a ideia de Colégio Invisível. Entendemos, assim, que o corpo de indivíduos atuante nas *digital humanities*, ao passo que reúnem as condições de uma comunidade de discurso reconhecidamente estabelecida e proeminente, também se caracteriza enquanto um Colégio Invisível, por estarem enquadrados na condição de "atividade bastante recente" e tendo nesta constatação sua principal preponderância. Para nós, portanto, a estrutura elitista das humanidades digitais corresponde ao colégio invisível da mesma por residir nela o principal núcleo de coletivos informais, pois ao considerarmos a interação entre indivíduos oriundos de diferentes áreas do conhecimento se daria sob uma ótica informal, com os agentes interagindo em meios que não são comuns ou próprios a todos estes.

A constatação de condições aplicáveis a um Colégio Invisível, formada a princípio por uma comunidade de discurso, nos enfatizou ainda mais se dedicar a análise das métricas de autor no estudo bibliométrico. As atividades cognitivas impressas nas atividades desses indivíduos que correspondem na prática, pelas práticas de citação,



forneem a matéria-prima para os estudos bibliométricos. Essas atividades compreendem fundamento para o estabelecimento de ligações entre autores e/ou seus trabalhos, revelando, porém, uma estrutura de um campo aplicada ao que se refere como 'indicador de proximidade cognitiva'. Portanto essa pesquisa se balizou pela premissa de que a vinculação de trabalhos e autores confere uma estrutura reveladora de um domínio.

A preocupação com a persecução da interpretação deste coletivo acadêmico apenas apresentaria condições de viabilidade se o mesmo partisse da literatura. Segue daí a importância da recuperação da literatura bem como a formatação da matriz de dados a serem analisados. Esse corpus de análise foi projetado com o devido rigor visando fornecer aquilo que de mais significativo a literatura pode representar. Ou seja, oferecer uma visão geral do estado corrente do desenvolvimento de um assunto, em um dado tempo, além de uma base conceitual, num processo unificador do conhecimento. Ainda também nos fornecendo uma orientação geral da nova área, já que a combinação de informação didática e bibliográfica torna as revisões um veículo ideal para fornecer informação básica para sobre determinado contexto acadêmico. Com efeito, esta pesquisa buscou através dessa revisão um estatuto teórico que possa valer como referencialidade diante das diversas posturas descritivas, metodológicas e técnicas diante do objeto, a partir dos textos que compõem a literatura.

Essa referencialidade necessitou de uma formatação em expressões de busca que fossem capazes de, ao mesmo tempo abranger um escopo de complexa mensuração, evitar que temáticas outras compusessem o corpus literário. Assim, houve o esforço para estabelecer uma abrangência controlada que nos permitiu abarcar não apenas a indexação explícita presente no título e/ou nas palavras-chave, além da estratégia de considerar o resumo com vistas a ampliar o conteúdo potencialmente revelador sobre a temática. De certo, consideramos que a estratégia de busca se verificou acertada, já que nos resultados revocados foi possível identificar conceitualmente a questão, todavia sem deixar de constar descobertas temáticas relacionadas. Da mesma maneira, acertada ainda, por não revocar quantidade significativa de informações ruidosas que fossem perceptíveis nesses resultados.

Ao partir para a análise efetiva do conteúdo recuperado, este estudo elegeu enquanto metodologia de trabalho a utilização do software adequado para análise das métricas da bibliometria e sua visualização. Dessa maneira nos foi possibilitado com a ajuda do VOSviewer promover, mesmo que em dimensão representativa, um mapeamento do campo, pois visualizando redes bibliométricas podemos inferir que

estamos diante de um "mapeamento da ciência", além de possibilitar a análise de centenas de artigos inviável a ação manual. A utilização do software nos serviu ainda enquanto validação da metodologia de pesquisa digital, na qual a ferramenta informática se responsabiliza pelo processamento do trabalho mais hercúlio, nos permitindo analisar mais dados e produzir mais conhecimento sobre os mesmos. Essa foi a garantia desta pesquisa em ter um cenário altamente representativo do domínio em questão.

No bojo dessas análises, o estudo produziu conhecimento novo a respeito de métricas amparadas nos três grupos estruturados pela pesquisa, sendo relativos aos autores, periódicos e palavras-chave. Nesse sentido analisamos a Citação de Referências, Periódicos, Autores, Organizações e Países. As cocitações foram baseadas nas análises de Referências citadas, Periódicos citados e Autores citados. As Coautorias para Autores além do Acoplamento bibliográfico de Referências, Periódicos, Autores e Organizações. Por fim, as Co-ocorrências de Palavras-chave.

No que tange a análise de citação de autores, periódicos e artigos/documentos podemos considerar, referindo às 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*) lidere exponencialmente as citações da área. Foi possível ainda identificar a participação da área da História no levantamento dessa métrica. Além disso, constatamos a relevância de um *cluster* consistente de autores da Ciência da Computação, tendo como marca principal um isolamento. De toda a forma evidenciamos a participação de autores que não formalmente ligados, em termos institucionais às *digital humanities*, como oriundos da Ciência da Informação e Computação. Ao que cabe à 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, que externam a proximidade entre autores, cuja proximidade é responsável por uma tecitura 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 grafos 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 à filosofia. 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 a 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 (*Information Studies*) assume um protagonismo ímpar por ditar com efusividade a temática das bibliotecas como as principais problematizações no domínio das *digital humanities*. 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 *digital humanities* 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.

O acoplamento bibliográfico foi observado pelo estudo enquanto conjunto de duas referências utilizadas pelos autores, evidenciando o ambiente intelectual no qual há uma relação cognoscível implícita. Esta característica pôde ser observada também em relação aos autores, documentos e periódicos. O *cluster* mais significativo do acoplamento de autores tem destaque para as autorias provenientes da área informacional já que autores da área da informação são, por exemplo, os maiores responsáveis por publicar e também aqueles com maior número de citações. Igualmente relevante que terceira autoria mais produtiva e citada desta métrica, desponte com ineditismo um ator que agrega à problematização das temáticas da altmetria – um temática própria dos estudos em Ciência da Informação. Nos prescindiui também sinalizar que três atores da área informacional possuem números de atividade bibliográfica muito superior, por exemplo, aos primeiros colocados ranqueados por força de *links*. A natureza do acoplamento de periódicos contribuiu para permitir uma identificação mais clara das áreas em evidência na literatura das humanidades digitais. 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. Diferentemente do caso do acoplamento de referências que serviu para exibir um cenário difuso, a acoplagem de periódicos complementou essa

abordagem e concedeu elementos suficientes para verificar a proeminência da área informacional na vetorização da literatura científica das humanidades digitais.

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. A análise foi baseada nas palavras-chave por atribuição pelas autorias, pelas bases de dados e, por fim, pela totalização de ambas. 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*

O cenário construído agora a partir dos termos atribuídos pelos autores, nos permitiu traçar as três categorias da mesma ordem que aplicamos enquanto categorização, sendo a primeira respectiva às áreas acadêmicas despontadas nesta observação. Dessa forma sugerimos *humanities, cultural heritage, history, digital history, education, humanities computing, digital scholarship e art*. Conforme objetos para as humanidades digitais identificamos *metadata, ontology, digital libraries, linked data, archives, digitization, technology, open access, crowdsourcing, rdf (Resource Description Framework), social network analysis, network analysis, digital curation, linked open data, digital preservation e data curation*. Para as práticas, ou sejam as atividades desenvolvidas nas HDs, sugerimos compreenderem *tei (text encoding initiatives), visualization, collaboration, digitization, text mining, crowdsourcing, data mining, network analysis, social network analysis, gis( geografic information systems) e text analysis*.

Categorizarmos ainda a observação destes termos em nível de indexadores atribuídos pela base de dados, apontamos enquanto áreas acadêmicas identificadas neste recorte: *education, history, social sciences, humanities, linguistic e information science*. Julgamos serem objetos para as humanidades digitais os termos: *digital libraries, metadata, information systems, linked data, information retrieval, social networking (online), ontology, big data, behavioral research, web services, human computer interaction, digital collections, internet, knowledge management, digital storage, information management, information technology, algorithms, user interfaces, research infrastructure, digital researches, hypertext systems e database systems*. Para as práticas de pesquisa consideramos que são apontadas pelos termos: *natural language processing systems, visualization, data mining, computational linguistics, data handling, artificial intelligence, learning systems, engineering education, virtual reality e hypertext systems*.

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 tecitura 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.

A tese de uma interlocução de áreas em níveis diferentes foi mote para este estudo. A identificação de termos e do corpo produtor de literatura das humanidades digitais possibilitou conhecer os detalhes do campo e verificar que algumas áreas contribuem efetivamente, corroborando, assim, essa tese. Todavia o estágio que este estudo alçou se propõe apenas afirmar que, devido a influência dessas áreas, será possível seguir contribuindo para heurística das *digital humanities*. Nesse sentido vemos um papel central nesse processo no qual áreas exercem despontada eloquência para os debates em HD. Por essa razão, defendemos nesta pesquisa que essa interlocução seguirá em evolução, diferenciando com mais propriedade os níveis de interdisciplinaridade no cerne do campo emergente. Nosso argumento se centrou na evolução tecnológica do ambiente informacional, base para as humanidades digitais, que passou por fortes transformações que não foram acompanhadas por todos os indivíduos no mesmo ritmo. Como as transformações pelo uso da tecnologia impactaram toda a atividade humana, a questão tem ainda mais relevância. Relevância essa ainda não identificada no campo das humanidades digitais, nem por sua representação temática, menos ainda por suas práticas.

Nesse contexto surgem as reflexões em torno da competência em informação que contempla esse processo em sua amplitude, todavia no âmbito da Ciência da Informação. Nosso ponto é que esta problematização se tornará emergente no escopo das humanidades digitais, tendo em vista sua predominante base de desenvolvimento/uso de ferramentas tecnológicas. Por essa razão julgamos imperativo conscientizar as reflexões em humanidades digitais dessa necessidade, haja vista sua característica primacial, que delega integralmente ao computacional a mediação entre informação e indivíduos. Nesse sentido vislumbramos nessa problemática mais um argumento para potencialização da congruência de escopos dos dois campos. Com efeito, ao identificar áreas em alta interlocução com as humanidades digitais, se compreendendo em interdisciplinaridade de diversos níveis, vemos nas problemáticas da competência em informação um promissor vetor para o crescimento da interdisciplinaridade com a Ciência da Informação.

Assim as *digital humanities* produzem, por sua natureza, uma convergência de reflexões que devem conduzir o campo para uma formatação própria dentro de um processo de consolidação típico da evolução de sua heurística. Não obstante é possível que a narrativa construída enquanto categorização deste movimento tenda a extenuar-se, consolidando os debates de ordem tecnológica não mais fora das áreas humanísticas em

particular, mas incorporando por estas uma naturalização temporal ainda vindoura. Mesmo consolidando quaisquer desses cenários, presumimos que a Ciência da Informação serviria como espaço qualificado para reflexões em e para as humanidades digitais tanto quanto suas condições de campo autônomo quanto narrativo além de, igualmente, agregá-lo teoricamente em ambas as situações.



## REFERÊNCIAS

- ALVARADO, Rubén Urbizagastegui. A lei de lotka na bibliometria brasileira. **Ciência da Informação**, v. 31, n. 2, p. 14-20, 2002. Disponível em: <http://www.brapci.inf.br/v/a/995>. Acesso em: 22 maio 2018.
- ALVES, Ana Paula Meneses; VIDOTTI, Silvana Aparecida Borsetti Gregório. O serviço de referência e informação digital. **Biblionline**, v. 2, n. 2, 2006. Disponível em <https://periodicos.ufpb.br/ojs2/index.php/biblio/article/view/611>. Acesso em: 16 out. 2018.
- ANDRIOLO, Eric. Desvendando 'Data Scraping': entenda como raspar dados pode facilitar o trabalho jornalístico. **Jornalismo nas Américas**, 2014. Disponível em: <https://knightcenter.utexas.edu/pt-br/blog/00-9586-desvendando-o-data-scraping-entenda-como-raspar-dados-pode-facilitar-o-trabalho-jornali>. Acesso em: 31 out.2017.
- ARAÚJO, Carlos Alberto. Bibliometria: evolução histórica e questões atuais. **Em Questão**, Porto Alegre, v. 12, n. 1, p. 11-32, jan/jun. 2006.
- BARONE, Fernando. A revisão de literatura em biblioteconomia e documentação: uma prática pedagógica. **Revista brasileira de Biblioteconomia e Documentação**, São Paulo, v. 23, n. 1/4, 1990, p. 136-146.
- BERRY, David M. (Org.). **Understanding digital humanities**. Nova Iorque: Palgrave Macmillan, 2012.
- BOREK, Luise; DOMBROWSKI, Quinn; PERKINS, Jody; SCHÖCH, Christof. TaDiRAH: a case study in pragmatic classification. **Digital Humanities Quarterly**, v. 10, n. 1, 2016. Disponível em <http://www.digitalhumanities.org/dhq/vol/10/1/000235/000235.html>. Acesso em: 18 abr. 2018.
- BOREK, Luise; DOMBROWSKI, Quinn; PERKINS, Jody; SCHÖCH, Christof. Building bridges to the future of a distributed network: from dirt categories to TaDiRAH, a methods taxonomy for digital humanities. *In: DCMI INTERNATIONAL CONFERENCE ON DUBLIN CORE AND METADATA APPLICATIONS*, 2014, Texas. **Electronic proceedings...** Disponível em: <https://dcpapers.dublincore.org/pubs/article/view/3720/1943>. Acesso em: 18 abr. 2018.
- BOYACK, K. W.; KLAVANS, R. Co-citation analysis, bibliographic coupling, and direct citation: Which citation approach represents the research front most accurately?. **Journal of the American Society for Information Science and Technology**, v. 61, n. 12, p. 2389–2404, dez. 2010.
- BRANDÃO, W. C.; PARREIRAS, F. S.; SILVA, A. B. O. E. Redes em Ciência da Informação: evidências comportamentais dos pesquisadores e tendências evolutivas das redes de coautoria. **Informação & Informação**, v. 12, n. 1, p. 110-124, 2007. DOI: 10.5433/1981-8920.2007v12n1espp110 Acesso em: 23 jul. 2019.

BROOKES, B. C. Bradford's law and the bibliography of science. **Nature**, n. 224, p. 953-956, 1969.

BRUCE, Christine Susan. Las siete caras de la alfabetización en información en la enseñanza superior. **Anales de Documentación**, n. 6, p. 289-294, 2003.

CAFÉ, Ligia Maria Arruda; BRÄSCHER, Marisa. Organização da informação e bibliometria. **Encontros Bibli: revista eletrônica de biblioteconomia e ciência da informação**, Florianópolis, p. 54-75, jan. 2008. ISSN 1518-2924. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2008v13nesp1p54>. Acesso em: 06 abr. 2018.

CAPURRO, Rafael et al. O conceito de informação. **Perspectivas em Ciência da Informação**, v. 12, n. 1, nov. 2007. Disponível em: <http://portaldeperiodicos.eci.ufmg.br/index.php/pci/article/view/54>. Acesso em: 05 abr. 2018.

CARIBÉ, Rita de Cássia do Vale. Comunicação científica: reflexões sobre o conceito. **Informação & Sociedade: Estudos**, v. 25, n. 3, 2015. Disponível em: <https://periodicos.ufpb.br/ojs2/index.php/ies/article/view/23109>. Acesso em: 11 maio 2018.

CASSUNDÉ, Fernanda Roda de Souza Araújo; BARBOSA, Milka Alves Correia; MENDONÇA, José Ricardo Costa. Entre revisões sistemáticas e bibliometrias: como tem sido mapeada a produção acadêmica em administração no Brasil?. **Informação & Informação**, v. 23, n. 1, p. 311-334, 2018. Disponível em: <http://www.brapci.inf.br/v/a/29715>. Acesso em: 11 maio 2018.

CASTELLS, M. **A Sociedade em rede**. São Paulo: Paz e Terra, 2011.

CHEN, Guo.; XIAO, Lu. Selecting publication keywords for domain analysis in bibliometrics: A comparison of three methods. **Journal of Informetrics**, v. 10, n. 1, p. 212-223, 2016.

COOPER, Harris; HEDGES, Larry V. Research synthesis as a scientific process. In: **The Handbook of Research Synthesis and Meta-Analysis**. New York: Russel Sage Foundation, 2009. Cap.1, p. 3-16.

COUTINHO, Eliana. Aplicação da lei de Bradford à literatura técnica sobre ferrovia: análise de periódicos e avaliação da base de dados da Rede Ferroviária Federal S.A. **Ciência da Informação**, v. 20, n. 2, p. 169-180, jul./dez. 1991.

CUNNINGHAM, Leigh. The librarian as digital humanist: the collaborative role of the research library in Digital Humanities Projects. **Faculty of Information Quarterly**, Toronto, v. 2, n. 2, Feb./Mar. 2010.

DAIGLE, Bradley J. The digital transformation of special collections. **Journal of Library Administration**, n. 52, p. 244-264, 2012.

DALBELLO, Marija. A genealogy of digital humanities. **Journal of Documentation**, v. 67, n. 3, p. 408-506, 2011.

DE BELLIS, Nicola. **Bibliometrics and citation analysis: from de Science Citation Index to cybermetrics**. Lanham, Maryland: Scarecrow Press, 2009.

DIAS, Célia da Consolação. A análise de domínio, as comunidades discursivas, a garantia de literatura e outras garantias. **Informação & Sociedade**, v. 25, n. 2, p. 7-17, 2015. Disponível em: <http://www.brapci.inf.br/v/a/18418>. Acesso em: 22 maio 2018.

VAN ECK, Nees Jan Van; WALTMAN, Ludo. Visualizing bibliometric networks. *In*: DING, Y.; ROUSSEAU, R.; WOLFRAM, D. (Ed.). **Measuring scholarly impact: methods and practice**. Leiden: Springer, 2014. p. 285-320.

VAN ECK, Nees Jan Van; WALTMAN, Ludo. **VOSviewer Manual**. Lieden: Universiteit Leiden, 2017. Disponível em: [http://www.vosviewer.com/documentation/Manual\\_VOSviewer\\_1.6.6.pdf](http://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.6.pdf). Acesso em: 10 jul. 2017.

VAN ECK, Nees Jan Van; WALTMAN, Ludo. **VOSviewer Manual: manual for VOSviewer version 1.6.6**. Leiden: University of Leiden, 2017. Disponível em: [http://www.vosviewer.com/documentation/Manual\\_VOSviewer\\_1.6.5.pdf](http://www.vosviewer.com/documentation/Manual_VOSviewer_1.6.5.pdf). Acesso em: 10 jul. 2017 .

EVENS, Aden. Web 2.0 and the Ontology of the Digital. **Digital Humanities Quarterly**, v. 6, n. 2, 2012. Disponível em: <http://www.digitalhumanities.org/dhq/vol/6/2/000120/000120.html>. Acesso em: 20 jul. 2018.

FALAGAS, M. E.; PITSOUNI, E. I.; MALIETZIS, G. A.; PAPPAS, G. Comparison of PubMed, Scopus, Web of Science, and Google Scholar: strengths and weaknesses. **The FASEB Journal**, v. 22, n. 2, p. 338–342, 2008.

FIGUEIREDO, Nice. Da importância dos artigos de revisão da literatura. **Revista brasileira de Biblioteconomia e Documentação**, São Paulo, v. 23, n. 1/4, p. 131-135, jan./dez. 1990.

FITZPATRICK, Kathleen. The Humanities, done digitally. *In*: GOLD, Matthew K. **Debates in the digital humanities**. Minneapolis: University of Minnesota Press, 2012, p. 3-67.

GARDINER, Eileen; MUSTO, Ronald G. **The digital humanities: a primer for students and scholars**. Cambridge: Cambridge University Press. 2015.

GARFIELD, Eugene. 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, 1984.

GARFIELD, Eugene. **Citation Indexing: its theory and application in Science, Technology and Humanities**. Philadelphia: ISI Press, 1983. Disponível em: <http://www.garfield.library.upenn.edu/ci/title.pdf>. Acesso em: 28 ago 2018.

GOLD, Matthew K. **Debates in the digital humanities**. Minneapolis: University of Minnesota Press, 2012.

GONZÁLEZ DE GOMÉZ, Maria Nélide A reinvenção contemporânea da informação: entre o material e o imaterial. **Pesquisa Brasileira em Ciência da Informação**, Brasília, v. 2, n. 1, p. 115-134, jan./dez. 2009.

GRÁCIO, Maria Cláudia Cabrini. Acoplamento bibliográfico e análise de cocitação: revisão teórico-conceitual. **Encontros Bibli**, v. 21, n. 47, p. 82-99, set. 2016. Disponível em: <https://periodicos.ufsc.br/index.php/eb/article/view/1518-2924.2016v21n47p82>. Acesso em: 05 abr. 2018.

GRÁCIO, Maria Cláudia Cabrini. Acoplamento bibliográfico e análise de cocitação: revisão teórico-conceitual. **Encontros Bibli**, v. 21, n. 47, p. 82-99, set./dez. 2016. DOI: 10.5007/1518-2924.2016v21n47p82.

GRÁCIO, Maria Cláudia Cabrini; OLIVEIRA, Ely Francina Tannuri de. Análise de cocitação de autores: um estudo teórico-metodológico dos indicadores de proximidade, aplicados ao GT7 da ANCIB. **Liinc em Revista**, v. 9, n. 1, 2013. Disponível em: <http://revista.ibict.br/liinc/article/view/3403>. Acesso em: 24 jul. 2019.

HARTSELL-GUNDY, Arianne; BRAUNSTEIN, Laura; GOLOMB, Liorah. **Digital humanities in the library: challenges and opportunities for subject specialists**. Chicago: Association of College and Research Libraries, a division of the American Library Association, 2015.

HJØRLAND, Birger. Citation analysis: a social and dynamic approach to knowledge organization. **Information Processing and Management**, v. 49, n. 6, p. 1313–1325, 2013.

HJØRLAND, Birger. Domain analysis in information science: Eleven approaches: traditional as well as innovative. **Journal of Documentation**, Londres, v. 58 n. 4, p. 422-462, 2002. <https://doi.org/10.1108/00220410210431136>.

HJØRLAND, Birger. Towards a theory of aboutness, subject, topicality, theme, domain, field, content ... and relevance. **Journal of the American Society for Information Science and Technology**, n. 52, p. 774-778, 2001.

HJØRLAND, Birger; ALBRECHTSEN, Hanne. Toward a new horizon in information science: domain-analysis. **Journal of the American Society for Information Science** n. 46, p. 400-425, 1995.

HUI, Yuk. What is a digital object?. **MetaPhilosophy**, v. 43, n. 4, p. 380-395, 2012.

INSTITUTO DE PESQUISA ECONÔMICA APLICADA. O que é? - Índice de Gini. **Desafios do Desenvolvimento**, v.1, n. 4, 2004. Disponível em:

[http://www.ipea.gov.br/desafios/index.php?option=com\\_content&id=2048:catid=28&Itemid=23](http://www.ipea.gov.br/desafios/index.php?option=com_content&id=2048:catid=28&Itemid=23). Acesso em: 25 maio. 2018.

KALLINIKOS, Jannis; AALTONEN, Aleks; MARTON, Attila. A theory of digital objects. **First Monday**, v. 15, n. 6, 2010. Disponível em: <http://firstmonday.org/ojs/index.php/fm/article/view/3033/2564>. Acesso em: 17 abr. 2018.

KIRSCHENBAUM, Matthew G. what is digital humanities and what's it doing in english departments?. **ADE Bulletin**, n. 150, 2010. Disponível em: [https://mkirschenbaum.files.wordpress.com/2011/01/kirschenbaum\\_ade150.pdf](https://mkirschenbaum.files.wordpress.com/2011/01/kirschenbaum_ade150.pdf). Acesso em: 01 dez 2015.

KIRSCHENBAUM, Matthew. The .txtual condition: digital humanities, born-digital archives, and the future literary. **Digital Humanities Quarterly**, v. 11, n. 3, 2013. Disponível em: <http://www.digitalhumanities.org/dhq/vol/7/1/000151/000151.html>. Acesso em: 19 jun 2018.

LEUNG, Xi. Y.; SUN, Jie; BAI, Billy. Bibliometrics of social media research: a citation and co-word analysis. **International Journal of Hospitality Management**, n.66, p.35–45, 2017.

LEYDESDORFF, Loet. Theories of citation?. **Scientometrics**. v. 43, n. 1, p. 5-25, 1998.

MANIFESTO das digital humanities. 2011. Disponível em: <https://tcp.hypotheses.org/497>. Acesso em: 26 de março de 2018.

MARQUES, Lilian Emanuelli ; PINHEIRO, Marta Macedo Kerr. A cúpula mundial sobre a sociedade da informação: foco nas políticas de informação. **Informação & Sociedade**, v.23, n.1, p. 117-131, jan./abr. 2013.

MCCARTY, W. “Humanities Computing.” *In: ENCYCLOPEDIA of Library and Information Science*. Nova Iorque: Marcel Dekker. 2003. Disponível em: <https://www.mccarty.org.uk/essays/McCarty,%20Humanities%20computing.pdf>. Acesso em: 04 de abril de 2019), p. 1224-1235.

MCCARTY, W. **Humanities computing**. Nova Iorque; Basingstoke: Palgrave Macmillan, 2005.

MEIRELES, M. R. G.; CENDÓN, B. V.; ALMEIDA, P. E. M. DE. Comparação do processo de categorização de documentos utilizando palavras-chave e citações em um domínio de conhecimento restrito. **Transinformação**, v. 28, n. 1, p. 87–96, 2016.

MIRANDA, Silvânia. Como as necessidades de informação podem se relacionar com as competências informacionais. **Ciência da Informação**, v. 35, n. 3, p. 99-114, set./dez. 2006.

MOREIRA, Walter. Os colégios virtuais e a nova configuração da comunicação científica. **Ciência da Informação**, Brasília, v. 34, n. 1, p. 57-63, jan. 2005.

Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0100-19652005000100007&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0100-19652005000100007&lng=en&nrm=iso). Acesso em: 03 ago. 2018.

MUELLER, Suzana Pinheiro Machado. O periódico científico. *In*: CAMPELLO, B. S.; CENDÓN, B. V.; KREMER, J. M. (Ed.). **Fontes de informação para pesquisadores e profissionais**. 2. reimpr ed. Belo Horizonte: Editora UFMG, 2007.

MUHANNA, Elias. What does “Born Digital” mean?. **International Journal of Middle East Studies**, v. 50, n. 1, p. 110-112, 2018.

NASCIMENTO, Denise Morado; MARTELETO, Regina. A “informação construída” nos meandros dos conceitos da teoria social de pierre bordieu. **DataGramZero**, v. 5, n. 5, 2004. Disponível em: <http://www.brapci.inf.br/v/a/2050>. Acesso em: 22 maio 2018.

NEWMAN, M. E. The structure of scientific collaboration networks. **Proceedings of the National Academy of Sciences of the United States of America** v. 98, n. 2, p. 404, 2001. Disponível em: <http://www.pnas.org/content/pnas/98/2/404.full.pdf>. Acesso em: 24 ago 2018.

OLIVEIRA, Ely Francina Tannuri de; ALVES, Bruno Henrique. Cosseno de Salton, Índice de Jaccard e Correlação de Pearson: comparando índices normalizados e absolutos em análise de cocitação de autores. **Em Questão**, Porto Alegre, v. 23, p. 235-253, Edição Especial 5 EBBC, 2017. Doi: <http://dx.doi.org/10.19132/1808-5245230.235-253>.

PINHEIRO, Lena Vania. Lei de bradford: uma reformulação conceitual. **Ciência da Informação**, v. 12, n. 2, p. 59-80, 1983. Disponível em: <http://www.brapci.inf.br/v/a/602>. Acesso em: 15 maio 2018.

PIMENTA, Ricardo Medeiros; GOMES, Josir Cardoso; GOUVEIA, Fabio Castro. GOOGLE SCHOLAR COMO FONTE DE MEDIÇÃO DA PRODUÇÃO CIENTÍFICA LUSÓFONA. 2018. *In*: CONGRESSO DE ALTMETRIA E CIÊNCIA ABERTA DA AMÉRICA LATINA, 1., 2018, Niterói. **Anais eletrônicos...** UFF: Niterói, 2018. Disponível em: <https://app.uff.br/riuff/handle/1/7928>. Acesso em: 3 fev. 2020.

POMBO, O. Interdisciplinaridade e integração dos saberes. **Liinc em Revista**, Rio de Janeiro, v. 1, n. 1, p. 3-15, 2005. Disponível em: <http://www.brapci.ufpr.br/brapci/v/a/5447>. Acesso em: 26 Out. 2017

PONOMARIOV, Branco; BOARDMAN, Craig. What is co-authorship?. **Scientometrics**, v. 109, n. 3, p. 1939–1963, 2016. Disponível em: <https://link.springer.com/content/pdf/10.1007%2Fs11192-016-2127-7.pdf>. Acesso em: 23 jul 2019.

PRICE, Derek John de Solla. **Little science, big science and beyond**. Columbia University Press: New York, 1986.

- PRICE, Derek John de Solla. **O desenvolvimento da ciência: análise histórica, filosófica, sociológica e econômica.** Rio de Janeiro: Livros Técnicos e Científicos, 1976.
- PROCÓPIO JÚNIOR, Pétersson S.; LAENDER, Alberto H. F.; MORO, , Mirella M. Análise da Rede de Coautoria do Simpósio Brasileiro de Bancos de Dados. *In:* SIMPÓSIO BRASILEIRO DE BANCO DE DADOS, 2011. **Anais eletrônicos...** Florianópolis, 2011. Disponível em: <https://pdfs.semanticscholar.org/4809/dd0708e90d29f30b005d69a81d5c3fd1bce3.pdf>. Acesso em: 24 ago. 2018.
- PROENÇA JÚNIOR, Domínio; SILVA, Édison Renato. **TransInformação**, Campinas, v. 28, n. 2, p. 233-240, maio/ago. 2016.
- QUONIAM, L. Bibliometrie sur des références bibliographiques: methodologie. *In:* DESVALS, H.; DOU, H. (Ed.). **La veille technologique.** Paris: Dunod, 1992. p. 244-262.
- QUONIAM, L. et al. Inteligência obtida pela aplicação de data mining em bases de teses francesas sobre o Brasil. **Ciência da Informação**, v. 30, n. 2, p. 20-28, 2001.
- RAO, I. K. Ravichandra. **Métodos quantitativos em Biblioteconomia e Ciência da Informação.** Brasília: Associação dos Bibliotecários do Distrito Federal: Organização dos Estados Americanos, 1986.
- ROSS, Lyman; SENNYEY, Pongracz. The Library is Dead, Long Live the Library!: the practice of academic librarianship and the digital revolution. **The Journal of Academic Librarianship**, v. 34, n. 2, p. 145–152, 2008.
- RUSSEL, Isabel Galina. ¿Qué son las Humanidades Digitales? **Revista Digital Universitaria**, v. 12, n. 7, jul. 2011. Disponível em: <http://www.revista.unam.mx/vol.12/num7/art68/index.html>. Acesso em: 21 mar 2018.
- SANTOS JUNIOR, Roberto Lopes dos. A abordagem teórica de Lena Vania Ribeiro Pinheiro sobre os conceitos inter e transdisciplinaridade. **TransInformação**, Campinas, v. 23, n. 3, p. 227-234, set./dez. 2011.
- SARACEVIC, Tefko. Ciência da informação: origem, evolução e relações. **Perspectivas em Ciência da Informação**, Belo Horizonte, v. 1, n. 1, p. 41-62, jan./jun. 1996.
- SCHREIBMAN, Susan. Digital humanities: centres and peripheries”. **Historical Social Research**, v. 37, p. 46-58, 2012.
- SCHREIBMAN, Susan; SIEMENS, Ray; UNSWORTH, Ray (Ed.). **A companion to digital humanities.** Oxford: Blackwell, 2004. Disponível em: <http://digitalhumanities.org/companion/>. Acesso em: 21 fev. 2018.
- SEGURA-MUÑOZ, S. I.; TAKAYANAGUI, A. M. M.; SANTOS, C. B.; SANCHEZSWEATMAN, O. Revisão sistemática de literatura e metanálise: noções

básicas sobre seu desenho, interpretação e aplicação na área da saúde. In: SIMPÓSIO BRASILEIRO DE COMUNICAÇÃO EM ENFERMAGEM, 8, 2002, SIBRACEN, Ribeirão Preto (SP). **Anais eletrônicos...** 2002. Disponível em: <http://www.proceedings.scielo.br/pdf/sibracen/n8v2/v2a074.pdf>. Acesso em: 19 fev. 2018.

SMALL, H. G.; KOENIG, M. E. D. Journal clustering using a bibliographic coupling method. **Information Processing & Management**, v. 13, n. 5, p. 277–288, 1977.

SMIRAGLIA, R. P. Two kinds of power: insight into the legacy of Patrick Wilson. **Annual Conference of CAIS / Actes du congrès annuel de l'ACSI**, 2013.

SOUSA, Marcos R. de; RIBEIRO, Antonio Luiz P. Revisão sistemática e meta-análise de estudos de diagnóstico e prognóstico: um tutorial. **Arquivo Brasileiro de Cardiologia**, São Paulo, v. 92, n. 3, p. 241-251, Mar. 2009. Disponível em: [http://www.scielo.br/scielo.php?script=sci\\_arttext&pid=S0066-782X2009000300013&lng=en&nrm=iso](http://www.scielo.br/scielo.php?script=sci_arttext&pid=S0066-782X2009000300013&lng=en&nrm=iso). Acesso em: 12 Set. 2018. <http://dx.doi.org/10.1590/S0066-782X2009000300013>.

SOUZA, Edivanio Duarte de. Configurações do campo da ciência da informação: pluralismo epistemológico e descentração interdisciplinar. **Tendências da Pesquisa Brasileira em Ciência da Informação**, v. 5, n. 1, não paginado, 2013. Disponível em: <https://revistas.ancib.org/index.php/tpbci/article/view/261/261>. Acesso em: 26 set. 2018.

SPINAK, Ernesto. **Diccionario enciclopédico de Bibliometria, cienciometria e informetria**. Caracas : Unesco -CII/II, 1996. Disponível em: <http://unesdoc.unesco.org/images/0024/002433/243329S.pdf>. Acesso em: 22 maio 2018.

SPIRO, Lisa. Getting Started in the Digital Humanities. **Digital Scholarship in the Humanities**, n. 14, out. 2011. Disponível em: <https://digitalscholarship.wordpress.com/2011/10/14/getting-started-in-the-digital-humanities/>. Acesso em: 20 abr. 2018.

SULA, Chris Alen. Digital humanities and libraries: a conceptual model. **Journal of Library Administration**, v. 53, n. 1, p.10–26, 2013.

SURWASE, G., et al. Co-citation analysis : an overview *In*: KADEMANI, B. S.; CENTRE FOR DEVELOPMENT OF ADVANCED COMPUTING (MUMBAI, INDIA); BOMBAY SCIENCE LIBRARIANS' ASSOCIATION (INDIA) (Ed.). **Beyond librarianship: creativity, innovation, and discovery**. Delhi: B.R. Pub. Corp, 2011.

SVENSSON, Patrik. Humanities Computing as Digital Humanities. **Digital Humanities Quarterly**. v. 3, n. 3, 2009. Disponível em: <http://ccftp.scu.edu.cn/Download/505594d1-330f-4831-94a1-ec87d218345a.pdf#page=176> Acesso em: 05 jan 2018.

TANG, Muh-Chyun; CHENG, Yun Jen; CHEN, Kuang Hua. A longitudinal study of intellectual cohesion in digital humanities using bibliometric analyses, **Scientometrics**, v. 113, n. 2, p. 985-1008 set. 2017. Disponível em: [https://www.researchgate.net/profile/Muh\\_Chyun\\_Tang/publication/319492607\\_A\\_lon](https://www.researchgate.net/profile/Muh_Chyun_Tang/publication/319492607_A_lon)



itudinal\_study\_of\_intellectual\_cohesion\_in\_digital\_humanities\_using\_bibliometric\_analyses/links/59b0c585a6fdcc3f8889b845/A-longitudinal-study-of-intellectual-cohesion-in-digital-humanities-using-bibliometric-analyses.pdf. Acesso em: 17 ago 2018.

TAXONOMY of Digital Research Activities in the Humanities - TaDiRAH. Disponível em: <http://tadirah.dariah.eu/vocab/index.php>. Acesso em: 17 abr. 2018.

TENNIS, J. T. Com o que uma análise de domínio se parece no tocante a sua forma, função e gênero?. **Brazilian Journal of Information Science**, v. 6, n. 1, 2012. Disponível em: <http://www.brapci.inf.br/v/a/13187>. Acesso em: 14 nov. 2017.

THATCAMP PARIS 2010, Paris, **Manifesto das digital humanities...** Paris, 2010. Disponível em: <http://tcp.hypotheses.org/category/manifeste>. Acesso em: 29 nov. 2018.

THIBODEAU, K. **Overview of technological approaches to digital preservation and challenges in coming yearsthe state of digital preservation: an international perspective**. Washington: CLIR and Library of Congress, 2002.

UNSWORTH, John. **Scholarly primitives: what methods do humanities researchers have in common, and how might our tools reflect this?**. London: King's College London. 2000. Disponível em: <http://people.virginia.edu/~jmu2m/Kings.5-00/primitives.html>. Acesso em: 17 abr. 2018.

VARELA, Aida Varela. 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...** 2005.

VANDEGRIFT, Micah; VARNER, Stewart. Evolving in common: creating mutually supportive relationships between libraries and the digital humanities. **Journal of Library Administration**, v. 53, n. 1, p. 67-78, 2013.

VANTI, Nadia Aurora Peres. Da bibliometria à webometria: uma exploração conceitual dos mecanismos utilizados para medir o registro da informação e a difusão do conhecimento. **Ciência da Informação**, v. 31, n. 2, p. 152-162, maio/ago. 2002.

VANZ, Samile Andréa de Souza; STUMPF, Ida Regina Chittó. Procedimentos e ferramentas aplicados aos estudos bibliométricos. **Informação & Sociedade**, João Pessoa, v. 20, n. 2, p. 67-75, maio/ago. 2010.

WEB OF SCIENCE. **Wikipedia, a enciclopédia livre**. Disponível em: [https://en.wikipedia.org/w/index.php?title=Web\\_of\\_Science&oldid=856456072](https://en.wikipedia.org/w/index.php?title=Web_of_Science&oldid=856456072). Acesso em: 7 set. 2018.

WEINBERG, B. H. Bibliographic coupling: a review. **Information Storage and Retrieval**, v. 10, n. 5-6, p. 189-196, May 1974.

YAMAOKA, Eloi Juniti; GAUTHIER, Fernando Ostuni. Objetos digitais: em busca da precisão conceitual. **Informação & Informação**, Londrina, v.18, n. 2, p.77-97, maio/ago. 2013.

ZHAO, Dangzhi; STROTMANN, Andreas. **Analysis and visualization of citation networks**. Vermont: Morgan & Claypool Publishers, 2015.

ZUCCALA, Alesia. Modeling the invisible college. **Journal of the American Society for Information Science and Technology**, n. 57, p. 152-168, 2006.