

TEC COGS

21

JAN./JUN.
2020

REVISTA DIGITAL DE
TECNOLOGIAS COGNITIVAS

ISSN
1984-3585

Humanidades Digitais

Programa de Pós-Graduação em
Tecnologias da Inteligência e Design Digital
Pontifícia Universidade Católica de São Paulo



Expediente

TECCOGS – Revista Digital de Tecnologias Cognitivas, n. 21, jan./jun. 2020, ISSN: 1984-3585
Programa de Pós-graduação em Tecnologias da Inteligência e Design Digital (TIDD),
Pontifícia Universidade Católica de São Paulo (PUC-SP)



Esta edição foi viabilizada por verba do Plano de Incentivo à Pesquisa (PIPEq) da PUC-SP.

DIRETOR CIENTÍFICO

Prof. Dr. Winfried Nöth
PUC-SP

VICE-DIRETORA CIENTÍFICA

Profa. Dra. Lucia Santaella
PUC-SP

EDITOR DO NÚMERO

Profa. Dra. Isabel Jungk

EDITOR EXECUTIVO

Prof. Dr. Guilherme Cestari

REVISÃO DE TEXTO E NORMATIZAÇÃO

Fábio de Paula
Marcelo de Mattos Salgado

CAPA E PROJETO GRÁFICO

Clayton Policarpo
Guilherme Cestari
Thiago Mittermayer

IMAGEM DA CAPA

Foto por [Vincent Eisfeld](#) em [Wikimedia](#)
Creative Commons ©

DIAGRAMAÇÃO E DIVULGAÇÃO ONLINE

Clayton Policarpo
Guilherme Cestari
Levy Henrique Bittencourt Neto
Thiago Mittermayer

CONSELHO EDITORIAL

Prof. Dr. Alex Primo
UFRGS
Prof. Dr. André Lemos
UFBA
Profa. Dra. Cláudia Giannetti
Barcelona
Profa. Dra. Clárisse Sieckenius de Souza
PUC-RIO
Profa. Dra. Diana Domingues
UnB FGA GAMA
Profa. Dra. Geane Alzamora
UFMG
Profa. Dra. Giselle Beiguelman
USP
Prof. Dr. João Teixeira
UFSCAR
Profa. Dra. Luiza Alonso
UnB
Profa. Dra. Maria Eunice Gonzales
UNESP-Marília
Prof. Dr. Ricardo Ribeiro Gudwin
UNICAMP
Prof. Dr. Sidarta Ribeiro
UFRN

n. 21, jan./jun. 2020

Sumário

Editorial Isabel Jungk	6
---------------------------	---

ENTREVISTA

Entrevista com Felipe Londoño Isabel Jungk	12
---	----

DOSSIÊ

Bases conceituais e tecnológicas para a modelagem de megadados em Humanidades Digitais Ana Maria Alfonso-Goldfarb, José Luiz Goldfarb, Márcia Helena Mendes Ferraz e Odécio Souza	25
---	----

ARTIGOS

A aliança entre humanidades e tecnologias computacionais e a resignificação de conhecimento Isabel Jungk	45
Um breve guia para as Humanidades Digitais Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner e Jeffrey Schnapp	69
Acesso aberto para as máquinas Christof Schöch	99
Researching visual semiotics online Jamin Pelkey	116
Comunicação secreta e história da criptologia Peter Krapp	146

EXTRA DOSSIÊ

A dialogue on cognitive semiotics: minds, and machines Vincent Colapietro, Winfried Nöth e Guilherme Cestari	167
Biopolítica digital e a utilização da inteligência artificial no controle de pandemias Paola Cantarini	185
Um debate sobre o campo online e a etnografia virtual Samira Cristina Silva Pereira e Sérgio Procópio Carmona Mendes	196

RESENHAS

- Resenha do livro *Os avanços das ciências podem acabar com a filosofia?*, de Ronaldo Marin e Gustavo Rick Amaral 214
Peter Schulz
- Resenha do livro *A new companion to digital humanities*, de Susan Schreibman, Ray Siemens e John Unsworth 218
Victor Sancassani
- Resenha do livro *Digital humanities: knowledge and critique in a digital age*, de David M. Berry e Anders Fagerjord 227
Mônica Allan

Editorial

Por Isabel Jungk¹

A Edição 21 da Revista TECCOGS tem o seu foco nas *Humanidades Digitais* (HD), o novo campo de pesquisa transdisciplinar que, desde os anos 2000, tem explorado as possibilidades de aplicar o potencial computacional em diversas disciplinas das ciências humanas e sociais.

As HD têm as suas origens nas áreas das tecnologias de visualização, da estruturação e da mineração de dados. Atualmente o campo das HD vem se expandindo, passando a abranger igualmente reflexões críticas sobre como os recentes avanços tecnológicos possibilitam novos caminhos para a pesquisa em humanidades, e como provocam impactos na produção e disseminação de conhecimento em todas as áreas do pensamento, com profundas consequências culturais e sociais.

Issue 21 of TECCOGS focuses on digital scholarship in the Humanities, the transdisciplinary research field that began to establish itself in the early 2000s and that has meanwhile consolidated itself as a new research field under the name of *digital humanities* (DH).

DH has its origins in the technologies of knowledge visualization, data mining, and big data structuring. Currently, the field of DH is expanding, to the study of the cultural and social consequences of data processing. Today, the DH also address critical reflections on how recent technological advances enable new venues for research in the humanities and how they influence the production and dissemination of knowledge in all areas of knowledge.

¹ Isabel Jungk é Doutora em Tecnologias da Inteligência e Design Digital e Mestre em Comunicação e Semiótica pela PUC-SP. cv Lattes: lattes.cnpq.br/2830878145479718. E-mail: isabeljungk@yahoo.com.br.

A presente edição estabelece um diálogo com pesquisadores da Colômbia, Estados Unidos, Canadá, Inglaterra, Alemanha, Noruega e do Brasil, reunindo artigos, entrevistas, traduções e resenhas que apresentam o estado da arte da pesquisa nas HD assim como estudos de caso.

Em entrevista concedida à TECCOGS, Felipe Cesar Londoño nos oferece uma verdadeira aula magna sobre as HD, compartilhando sua longa experiência em projetos desenvolvidos na área de design e artes midiáticas, bem como sua visão crítica das potencialidades das HD para a transformação da universidade e da sociedade.

Ana Maria Alfonso-Goldfarb, José Luiz Goldfarb, Márcia Helena Mendes Ferraz e Odécio Souza apresentam o projeto desenvolvido no Centro Simão Mathias (PUC-SP) envolvendo cooperação internacional a fim de tornar mais eficiente a pesquisa e o trabalho documental em História da Ciência por meio da elaboração de ferramentas para a modelagem de grandes bases de dados, tendo em vista sua indexação e classificação.

Isabel Jungk apresenta a possibilidade de novas sínteses informacionais resultantes da conjugação das tecnologias computacionais e da pesquisa em humanidades, processo que ressignifica o conhecimento passado e abre perspectivas futuras para pensar criticamente a produção científica e cultural.

This edition establishes a dialogue with scholars in the field of the DH from Colombia, the United States, Canada, England, Germany, Norway, and Brazil. In articles, interviews, translations and reviews it presents the state of the art in DH as well as exemplary case studies.

In an interview with TECCOGS, Felipe Cesar Londoño presents a true inaugural lecture on DH, sharing his long-standing Colombian experience in projects developed in the area of design and media arts, together with his critical view concerning the potential of DH for the transformation of the university and the society.

Ana Maria Alfonso-Goldfarb, José Luiz Goldfarb, Márcia Helena Mendes Ferraz, and Odécio Souza present the project developed at Simão Mathias Center of PUC-SP and its international cooperative network. Exploring the potential of digital technologies for research and documentation in the History of Science, they show how research becomes more effective through the development of tools for modeling large databases with a view to their indexing and classification.

Isabel Jungk discusses the possibility of new informational syntheses resulting from the combination of computational technologies and research in the humanities, a process that offers new meaning to past knowledge and opens future perspectives to think critically the scientific and cultural production.

Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner e Jeffrey Schnapp trazem um guia prático para as HD, apresentado os fundamentos e características do campo. Os autores traçam um roteiro dos passos relevantes para quem busca desenvolver projetos em HD, discutindo o projeto como unidade básica, o papel das instituições que os promovem, e a forma de avaliar os resultados.

Christof Schöch discute a necessidade de modelos e soluções para que as publicações científicas possam ir além do formato pdf como reprodução digital do texto impresso, a fim de serem disponibilizadas como dados abertos semiestruturados e vinculados, possibilitando sua legibilidade pelas máquinas.

Jamin Pelkey aborda a pesquisa em dados visuais e a análise de seus significados a partir de estudos de caso, mostrando o potencial da teoria semiótica para a formação de hipóteses, parâmetros de pesquisa e de codificação, bem como para a construção de bases de dados visuais complexos por meio de ferramentas de busca no contexto das HD.

Peter Krapp traça uma perspectiva histórica da necessidade de sigilo nas comunicações através da criptologia, atualmente fundamentada na matemática e computação, como contraponto ao compartilhamento na era da cultura digital, e como isso representa um desafio para as HD.

Anne Burdick, Johanna Drucker, Peter Lunenfeld, Todd Presner, and Jeffrey Schnapp present a practical guide for DH together with a panorama of the fundamentals and characteristics of the field. The authors outline a script of the relevant steps for those looking to develop DH projects, discussing the project as a basic unit, the role of the institutions that promote them, and the way to evaluate the results.

Christof Schöch discusses the necessity of models and solutions for scientific publications to advance beyond the pdf format as a digital reproduction of printed text, and to make digital publications available as open semi-structured and linked data that enables open access as well as their readability by machines.

Jamin Pelkey's topic is research in visual data and the analysis of their meaning on the basis of case studies, showing the potential of semiotic theory for the formation of hypotheses, coding and research parameters as well as for the construction of complex visual databases and its collection through search engines in the context of DH.

Peter Krapp traces a historical perspective of the need for secrecy in communications through cryptology, currently grounded in mathematics and computing, as a counterpoint to sharing in the era of digital culture and how it represents a challenge for DH.

Pela primeira vez, a revista TECCOGS publica artigos além do dossiê temático. No primeiro deles, esta edição traz um diálogo entre Vincent Colapietro e Winfried Nöth sobre o campo da semiótica cognitiva e sobre artefatos tecnológicos como extensões da mente humana.

Nos artigos seguintes, Paola Cantarini apresenta as relações entre inteligência artificial e biopolítica digital por meio de uma abordagem interdisciplinar e crítica sobre o controle realizado via manipulação de dados no atual contexto pandêmico, e Samira Pereira e Sérgio Mendes debatem o potencial e os desafios da etnografia digital como ferramenta de pesquisa antropológica no contexto da sociedade da informação, junto com suas inovações metodológicas e capacidade para estabelecer novas relações entre os pesquisadores e seus interlocutores.

A seção de resenhas apresenta três livros de interesse para a interdisciplinaridade e as pesquisas em HD. Peter Schulz aborda o livro *Os avanços das ciências podem acabar com a filosofia?* de Ronaldo Marin e Gustavo Rick Amaral, que discute as relações entre ciência e filosofia a partir da perspectiva da unificação conceitual.

Victor Sancassani apresenta o livro *A new companion to digital humanities*, editado por Susan Schreibman, Ray Siemens e John Unsworth, versão atualizada do livro que marcou o início do movimento das HD em 2001, e que traz uma fértil coleção de artigos cobrindo os mais variados aspectos das HD.

For the first time, TECCOGS publishes nonthematic articles. In the first one, this issue brings a dialogue between Vincent Colapietro and Winfried Nöth on cognitive semiotics and the technological extensions of the human mind.

In the following articles, Paola Cantarini discusses the relationship between artificial intelligence and digital biopolitics. Adopting an interdisciplinary and critical approach, she studies the methods of control performed via data manipulation in the pandemic context. Samira Pereira and Sérgio Mendes inquire into the potential and the challenges of digital ethnography as an anthropological research tool in the context of information society, together with its methodological innovations and the ability to establish new relationships between researchers and their interlocutors.

The reviews section presents three books of interest to interdisciplinarity and research in DH. Peter Schulz reviews the book *Os avanços das ciências podem acabar com a filosofia?*, by Ronaldo Marin and Gustavo Rick Amaral, which discusses the relationship between science and philosophy based on the concept of conceptual unification.

Victor Sancassani presents the book *A new companion to digital humanities*, edited by Susan Schreibman, Ray Siemens and John Unsworth, an updated version of the book that marked the beginning of the DH movement in 2001, which brings a fertile collection of articles covering the various aspects of DH.

Ao final, Monica Allan apresenta o livro *Digital humanities: knowledge and critique in a digital age*, de David Berry e Anders Fagerjord, que busca fornecer os elementos para uma crítica fundamentada da produção de conhecimento na era digital.

Somos gratos aos que a tornaram possível esta nova edição com suas contribuições.

Finally, Monica Allan presents the book *Digital humanities: knowledge and critique in a digital age*, by David Berry and Anders Fagerjord, which seeks to provide the elements for a well-founded critique of knowledge production in the digital age.

We are grateful to all those whose contributions have made this new edition possible.



ENTREVISTA

Entrevista com Felipe Londoño

Por Isabel Jungk¹

Felipe César Londoño López é arquiteto pela Universidade Nacional da Colômbia (Sede Manizales) e doutor em Engenharia Multimídia pela Universidade Politécnica da Catalunha, Espanha. Em sua longa experiência profissional, desenvolveu atividades de ensino, pesquisa, desenvolvimento e administração em diversas instituições.

Foi Professor Associado da Universidade de Caldas, onde foi Reitor e cofundador do Departamento de Design Visual. Professor nas áreas de design visual experimental e pesquisador de arte, design e novas mídias, é diretor do *Festival Internacional da Imagem*, evento especializado em artes eletrônicas e cultura digital. Também desenvolveu projetos junto aos Ministérios da Cultura e da Educação da Colômbia. Ele é Reitor da Faculdade de Artes e Design da Universidade Jorge Tadeo Lozano de Bogotá e foi professor participante do Doutorado em Engenharia Multimídia na Universidade Politécnica da Catalunha, Espanha. Recebeu diversos prêmios e menções honrosas em suas áreas de atuação.

Ao longo de sua carreira acadêmica, ele se especializou nas relações entre interface, pessoas e mídias. Pesquisador de novas tecnologias e seus impactos, publicou livros e artigos sobre linguagens e arte digital, design visual e espaços públicos. Sua pesquisa aborda temas como expressão visual (1994), ambientes visuais e virtuais (2011), padrões de cores e arquitetura (2006), artes midiáticas, cinema, videogames e cidadania (2012), interfaces e comunidades virtuais (2004), design digital (2009) e suas metodologias em projetos interativos (2006), métodos didáticos, práticas curatoriais e empreendedorismo cultural, entre outros.

Membro de vários conselhos editoriais, organizou coleções de livros e, recentemente, o livro *Dimensões: arte_design_tecnologia* (Goiânia: CEGRAF/UFG, 2020), junto a pesquisadores brasileiros. Tanto na Universidade de Caldas como na Universidade Jorge Tadeo Lozano, tem trabalhado continuamente na criação de programas de design e

¹ Isabel Jungk é Doutora em Tecnologias da Inteligência e Design Digital (TIDD) e Mestre em Comunicação e Semiótica pela PUC-SP. cv Lattes: lattes.cnpq.br/2830878145479718. E-mail: isabeljungk@yahoo.com.br.

arte midiática, promovendo alianças internacionais, gerando intercâmbios com pesquisadores de diferentes países para fortalecer os processos de criação, design e arte e sua integração com tecnologias e outras áreas da ciência.

Ao despedir-se do cargo de Reitor da Universidade de Caldas, Londoño (2018) afirmou sua convicção vigorosa “na necessidade de transformar as visões da Universidade, fruto de um esforço que se torna cada vez mais vital como projeto necessariamente coletivo”, o que considera um dever que ratifica sua “vocação para uma Universidade aberta, crítica, comprometida com o público e em permanente construção coletiva”.

Esta entrevista com o Professor Londoño, concedida à revista TECCOGS, procura explorar o tema das interconexões entre as novas tecnologias e as humanidades de uma forma ampla, e mais especificamente o jovem movimento das Humanidades Digitais, contando com a sua rica experiência e visão acadêmicas para iluminar os caminhos do conhecimento e do futuro eu desejamos construir.

I.J.: Em 2019, em Manizales, foi organizado o 1º Colóquio Internacional de Humanidades Digitais, resultado da cooperação acadêmica entre Colômbia, França e Peru, que teve como objetivo promover a formação e a pesquisa em Humanidades Digitais. O nome, Humanidades Digitais (HD), é amplamente utilizado e designa um campo aberto de pesquisas em desenvolvimento. Na sua perspectiva, o que são as HD hoje? Como elas se distinguem das humanidades tradicionais?

F.L.: É possível afirmar que o conceito de Humanidades Digitais se refere ao uso de ferramentas tecnológicas de informação e comunicação, aplicadas às disciplinas humanísticas, seja nos processos de ensino e pesquisa, ou de criação e de planejamento social. No entanto, essa definição pode ser muito curta para estabelecer o verdadeiro escopo do que hoje é considerada uma revolução no pensamento humano, a partir das tecnologias digitais.

Os processos de troca de informações do ser humano, nos quais os dados convergem com as tecnologias interativas, constituem mais uma etapa na evolução do conceito de comunicação humana, como em sua época o foram a imprensa, a fotografia ou o cinema. Eles fazem parte essencial na consolidação de uma sociedade cada vez mais permeada pelo universo visual e pelas redes de informação. As novas tecnologias de informação e comunicação têm gerado um crescimento acelerado dos

conteúdos disponibilizados em redes e aplicativos digitais, paralelamente às possibilidades de integração dos diferentes meios e ao incremento das funcionalidades e benefícios baseados na interatividade dos aplicativos. Mas isso não é novo.

A busca por essa interrelação entre informação e tecnologia havia começado antes com um pequeno ensaio intitulado “As We May Think”, escrito por Vannevar Bush e publicado na edição de julho de 1945 da revista *Atlantic Monthly*. Bush, um cientista militar de alto escalão, descreve um processador de informações teóricas chamado Memex, que permitia ao usuário dispor de uma grande quantidade de informações, semelhante aos conhecidos navegadores da Internet. Como Bush, pioneiros como Ivan Sutherland, Douglas Engelbart, Marshall McLuhan, Ted Nelson e muitos outros, estabeleceram as bases para pensar sobre tecnologia e dispositivos, para além de simples ferramentas de conexão. É sob essas premissas que se propõe, a partir da Universidade de Caldas, em aliança com a Universidade Paul Valéry/Montpellier III, na França, e a Universidade de Piura, no Peru, a criação de cenários de discussão sobre as Humanidades Digitais, no marco da convocatória do PREFALC 2017 – Programa Regional França/América Latina e Caribe – e com o apoio da Fundação Maison des Science de L’homme (FMSH) (Paris).

Esta aliança motivou a realização do Primeiro Colóquio Internacional de Humanidades Digitais em Manizales em 2019 e com ele, a proposta de um Mestrado em Humanidades Digitais que investiga o que hoje significa compreender as Humanidades no contexto da acessibilidade e socialização de conhecimentos, da criação interdisciplinar, intercultural e colaborativa, a partir do pensamento crítico e da transferência dos meios digitais para o bem-estar humano.

I.J.: Como os modos de pesquisa com meios tecnológicos transformam a produção de conhecimento nas disciplinas humanísticas?

F.L.: As Humanidades Digitais recuperam diversos dispositivos que são aplicados em processos de investigação nos âmbitos da criação e das ciências sociais e humanas. A intersecção entre dados, informações e tecnologias possibilita promover a ciência aberta, a análise de dados aplicada à gestão da informação georreferenciada, entre tantas outras possibilidades, o que amplia a escala de ação e compreensão e possibilitam as bibliotecas digitalizadas que facilitam cartografias, arquivos e repositórios culturais, além de campi virtuais para processos de formação ou

intercambio de conhecimento. A tecnologia digital, como ponte entre a informação, o conhecimento e as comunidades, enfrenta pelo menos esses três grandes desafios: do ponto de vista instrumental, os dispositivos e os softwares possibilitam processos ágeis de comunicação e apropriação social do conhecimento. Graças a eles, os resultados de pesquisas especializadas chegam às pessoas de forma muito mais direta e podem ser por elas apropriados em seu próprio benefício.

Por outro lado, a pesquisa com meios tecnológicos expande seus métodos para processos que antes não eram considerados. A análise algorítmica de sistemas processuais se tornou possível graças ao Big Data e aos complexos sistemas de Inteligência Artificial, que facilitam a observação de problemáticas complexas, ligadas a questões de saúde pública (os estudos sobre a atual pandemia são evidências disso), do meio ambiente, de temas de justiça territorial e da paz.

Mas também a tecnologia ligada à pesquisa nos convida à criação coletiva (cocriação), ao estabelecimento de redes e ao desenvolvimento de processos colaborativos que, embora não sejam novos, se potencializam com o surgimento dos meios digitais. Plataformas como Leonardo / The International Society for the Arts, Sciences and Technology (Leonardo/ISAST), o MediaLab, Laboratório de Pesquisa, Desenvolvimento e Inovação em Mídias Interativas da UFG (Universidade Federal de Goiás), o TadeoLab da Universidade de Bogotá Jorge Tadeo Lozano ou o Centro de Ciências Francisco José de Caldas da Universidade de Caldas, na Colômbia, promovem espaços de construção e participação cidadã com programas que integram arte e ciência com tecnologia, design e inovação para gerar soluções para problemas complexos com ferramentas digitais colaborativas e de cocriação.

I.J.: Como se modifica o ensino e a transmissão do conhecimento na universidade a partir das novas possibilidades tecnológicas? Que mudanças já podem ser observadas na Colômbia, no Brasil e em outros países?

F.L.: Num contexto amplo, é possível falar de uma evolução do conceito de Universidade, especialmente agora que se apresentam estas novas alternativas tecnológicas. Teóricos como Immanuel Wallerstein (2006) mencionam como as Instituições de Ensino Superior passaram por profundas mudanças, pelo menos em três momentos da história. A primeira delas aconteceu por volta do ano 1100, quando as Universidades foram criadas para proteger os conhecimentos e a cultura, que corriam

risco com as invasões da época. Um segundo momento é definido por uma Universidade que se integra na dinâmica da era industrial. Esta é, segundo Wallerstein, a universidade como a conhecemos hoje, com disciplinas que respondem às necessidades de desenvolvimento da indústria e do comércio. Por fim, é nas últimas duas décadas que surge uma Universidade diferente, que alguns chamam de “Terceira Geração”, e que se caracteriza pela utilização das tecnologias para se ligar, de forma mais direta, à sociedade. É uma pluriversidade, que se amplia por meio das redes e avanços nas tecnologias da informação e comunicação. Em suma, é uma instituição que se transforma, não apenas pelas novas formas de produção de conhecimento, mas principalmente pela velocidade com que esse conhecimento é intercambiado, apropriado e aplicado.

Experiências recentes, que temos compartilhado entre Universidades de vários países latino-americanos, demonstram isso. Assistimos à criação de plataformas conjuntas de intercâmbio de saberes, graças às redes que nos permitem o acesso à informação e aos dados. Assim, nossas instituições universitárias se transformam em cenários de diálogo onde o local é valorizado, em diálogo com o mundo.

Na prática, criamos *Makerspace* (como o Centro de Ciências Francisco José de Caldas na Colômbia), *MediaLabs* (como os da UFG e UCaldas), cursos de pós-graduação online conjuntos (como o Curso de Pós-graduação Online em Artes Mediáticas em que participam a Universidade do Chile, a Universidade de Córdoba na Argentina e a Universidade de Caldas), cátedras compartilhadas (como as “Conversas do Fim do Mundo” com a Universidade do Chile, a Universidade de Buenos Aires, a UFG do Brasil e a Universidade Jorge Tadeo Lozano da Colômbia), e publicações em rede que conectam conhecimento próprios com os saberes globais. Questões como interculturalidade, empreendedorismo social, responsabilidade social e ambiental, inovação pública, ciência aberta, letramento digital etc., adquirem um valor para além da produção tradicional de conhecimento. E muito disso se deve ao uso da tecnologia digital que se torna fundamental para a relação com a sociedade e seus territórios.

I.J.: As HD aproximam as disciplinas humanísticas das científicas para pensar os desafios humanos?

F.L.: A cultura digital conseguiu transformar o que C. P. Snow chamou de separação das “duas culturas”. As tecnologias digitais, entre outros fatores, têm levado artistas e cientistas, que durante décadas trabalharam separadamente, a se unirem para pensar os problemas mundiais. As

Humanidades Digitais promovem uma aproximação entre arte e ciência, o que fica evidente na forma como as duas utilizam instrumentos comuns e metáforas semelhantes para enfrentar os grandes desafios.

Temas como sustentabilidade ambiental ou crise energética são analisados hoje por equipes de cientistas, pesquisadores e artistas, para promover ações alternativas que possibilitem medidas de mitigação de riscos. Paralelamente aos sistemas regulatórios da ciência e tecnologia, essas comunidades desenvolvem trabalhos em grupos interdisciplinares abertos e com processos de participação cidadã que promovem o desenvolvimento de infraestruturas colaborativas para a solução de problemas sociais e ambientais, para além dos interesses do mercado global. A razão é simples: perante os problemas complexos que nos ameaçam, é necessário recorrer a outras visões que tradicionalmente não foram levadas em conta, de forma que essas múltiplas perspectivas participem na busca de soluções concretas. E acontece ao pensar a ciência de forma aberta e participativa, que possibilite a participação e administração dos bens comuns, que recorra às redes eletrônicas, ao software livre e ao livre acesso de publicações e dados, antes restritos a poderes hegemônicos.

Se a grande maioria das informações sobre o mundo é mediada por instrumentos científicos, as Humanidades Digitais fornecem uma visão “sensível” da relação natureza-cultura através de perspectivas interdisciplinares, que sugerem soluções alternativas para problemas locais e globais, com vistas à criação de condições justas para um desenvolvimento verdadeiramente sustentável.

I.J.: Diante dos avanços científicos alcançados, como as HD podem estimular a valorização e o desenvolvimento das disciplinas humanísticas neste contexto?

F.L.: Sem dúvida, os avanços da ciência favoreceram o fortalecimento das Humanidades Digitais. Em sentido inverso, as Humanidades Digitais também têm beneficiado, com suas novas perspectivas, os múltiplos campos do conhecimento das ciências nas esferas biomédica, bioinformática, nas ciências exatas e naturais, nas ciências agropecuárias, entre outras.

Mas, por outro lado, a transdisciplinaridade das Humanidades Digitais potencializa as disciplinas humanísticas tradicionais, ao enriquecê-las com métodos típicos das ciências, vinculados às esferas sociais. Processos relacionados ao código aberto, mapeamento, design de jogos, realidade

aumentada, análise algorítmica, conhecimento ubíquo, mineração de dados, computação em nuvem, prosa multimídia, cenários transmídia, robótica e inteligência artificial, desterritorializam a prática humanística e promovem a criação participativa de conteúdos abertos, as redes colaborativas, as cartografias digitais, a análise dos territórios inteligentes, as economias criativas, conectografias rurais, o design e o planejamento, entre outros.

I.J.: Pela sua experiência como diretor do Festival Internacional da Imagem da Universidade de Caldas, cujo tema em 2019 foi o Humanismo Digital, como os novos meios e linguagens tecnológicas vem impactando os processos criativos em geral, e a arte e o design especificamente?

F.L.: O tema central do 18º Festival Internacional da Imagem, em 2019, foi o Humanismo Digital, e isso aludiu aos processos de interação do ser humano com as tecnologias digitais e as possibilidades que esses processos abriram diante de novos mecanismos de participação cidadã para o diálogo, o intercâmbio cultural e criativo, a cocriação artística contemporânea, a produção de conhecimento aberto e a visibilidade das comunidades criativas. Falar de humanismo digital, no contexto do Festival da Imagem, implicou aprofundar a necessidade de incluir, não só a tecnologia nos processos de criação, mas também o pensamento crítico como eixo transversal necessário à análise ao aproveitamento da informação e do conhecimento.

As diferentes experiências apresentadas durante o evento evidenciarão uma transformação nas linguagens criativas e nos processos experimentais com as tecnologias, graças às dinâmicas observadas a partir de práticas transdisciplinares em laboratórios abertos, no contexto de visões *bottom-up*, *open source*, design participativo em comunidades, coletivos emergentes, hibridizações conceituais, tecnologias transparentes e arquivamento e preservação digital. Em suma, uma nova gramática que enriquece as linguagens da criação nas artes e no design e que pôde ser observada através de obras como as dos artistas colombianos Juan Cortés e do coletivo Atractor (“Colisión”, que se propõe assumir a noção de colisão do ponto de vista histórico do continente); Alejandro Londoño Urdaneta (que apresenta uma obra onde as máquinas informáticas parecem se apropriar do mundo; Estefanía García Pineda (“Rumo a uma geografia corporal do conflito”, cujo eixo central é a memória coletiva de sua família); Juan Pablo Pacheco (que, em “Liquid Democracy”, imagina sistemas digitais de participação cidadã), ou Juan José García (que indaga sobre uma visão da Colômbia no futuro, a partir de experiências com o projeto Mars Desert Research Station).

I.J.: *Como as tecnologias da imagem e os novos meios de visualização mudam as relações do grande público com o conhecimento e a cultura?*

F.L.: Em tempos como os que vivemos, as tecnologias da imagem e os dispositivos de visualização de dados são fundamentais para o desenvolvimento de estratégias de apropriação social do conhecimento e da cultura. Essa apropriação social é um processo que convida à compreensão dos fenômenos através das relações entre arte, ciência e tecnologia, e isso contribui, não só para o fortalecimento da cultura científica, mas também incentiva a participação ativa de diferentes pessoas, todas aquelas que tenham a capacidade de gerar conhecimento e, a partir da cultura como eixo fundamental, se apropriar da tradição e da inovação com uma visão global a partir do local.

A troca de conhecimentos, através da concepção e implementação de metodologias que promovam o diálogo entre diferentes pessoas e comunidades, é fundamental para o fortalecimento de capacidades nos contextos locais. Para que isso se concretize, será necessário gerar espaços de apropriação social do conhecimento que promovam a cultura e a participação das comunidades, por meio dos quais o cidadão se sinta cogestor e corresponsável pelos desenvolvimentos científico-tecnológicos que possam ter implicações e impactos em seu contexto local e regional.

I.J.: *O que é importante ter em mente quando se deseja fazer pesquisa na área de humanidades utilizando meios tecnológicos? Na sua opinião, quais são as principais interconexões entre as disciplinas humanísticas e as possibilidades tecnológicas?*

F.L.: A tecnologia é mais do que o dispositivo que a suporta. A tecnologia contribui para modular um conhecimento e essa avaliação é fundamental para entender a importância das Humanidades Digitais. Para além do meramente instrumental, a tecnologia digital permite o pensamento crítico e a criatividade no quadro do que hoje conhecemos como “sociedade do conhecimento”. Arnold Pacey afirma que é necessário valorizar a tecnologia como atividade humana e como parte da vida (“A Cultura como Tecnologia”, 1990), na medida em que contém aspectos que podem estar ligados a dimensões culturais. Como prática social, afirma Pacey, a tecnologia aborda aspectos técnicos (relacionados a capacidades, habilidades técnicas, matérias-primas, produtos, entre outros), dimensões organizacionais (como gestão política e administrativa) e temas ideológico-culturais que se relacionam com sistemas de valores e códigos éticos universais.

A evolução tecnológica, portanto, transforma a sociedade atual e possibilita desenvolvimentos inéditos, o que gera uma interconexão com as disciplinas humanísticas que deve ser levada em consideração. Na perspectiva da educação, por exemplo, o desafio é grande e implica a transformação de um sistema que está há anos numa tradição que importa rever, para estar em consonância com as expectativas de uma geração que cresceu entre telas e videogames. A hipermedialidade, graças às tecnologias da informação, permite que os conteúdos acadêmicos integrem texto, imagem, som, vídeo e animações para, com isso, proporcionar experiências amplas e profundas aos alunos, em espaços de informação que se interconectam com redes complexas que ampliam o conhecimento de cada pessoa para fortalecer seu conhecimento e cultura.

As Humanidades Digitais também possibilitam uma visão crítica do uso e apropriação tecnológica, a partir da qual é importante exigir e observar como as tecnologias digitais devem ser cada vez mais adaptativas e adequadas a cada contexto específico onde são aplicadas e buscam alternativas para gerar autonomia e resistência às hegemonias que promovem a dependência comercial e promover a cultura livre para incentivo e o letramento digital que viabilizam o bem comum.

I.J.: Como as mídias digitais e suas virtualidades podem ser utilizadas para impactar e desenvolver o cenário social e seus problemas (ambientais, econômicos, políticos etc.)?

F.L.: Eu acredito que são muitas as formas pelas quais as mídias digitais abrem o caminho para a participação e a solução dos problemas atuais. As redes, utilizadas de maneira inteligente, dão voz a muitas comunidades que tradicionalmente não têm conseguido expressar suas opiniões sobre temas de interesse comum e encontrar canais de comunicação alternativos de visibilidade e busca de soluções aos problemas e conflitos dos grupos humanos. Graças às tecnologias, é possível compreender a participação cidadã para além das narrativas modernistas e, por meio dos fluxos de informação que transitam em suas redes, é possível detectar movimentos sociais e manifestações políticas, bem como comunidades virtuais e práticas que apresentam outras formas de compreender os mundos que habitamos. Por outro lado, graças a trabalhos como os realizados pela artista colombiana Bárbara Santos, é possível explorar o hábitat, os costumes e os conhecimentos dos sábios indígenas que vivem na Amazônia colombiana. Utilizando técnicas como a realidade aumentada, Bárbara nos leva a universos ancestrais e nos apresenta um “outro” mundo onde o respeito pela natureza é a base da convivência.

Por outro lado, Roger Malina, doutor em Física, astrônomo e destacado professor da Universidade do Texas, promove o uso dos meios digitais para estabelecer a relação entre arte e ciência. Malina afirma que a grande maioria das informações sobre o mundo é mediada por instrumentos científicos, o que impede uma aproximação real a muitos dos problemas de hoje, e propõe valorizar uma visão “sensível” da natureza, por meio de perspectivas interdisciplinares que sugerem soluções alternativas ao controle das mudanças climáticas, da produção de outras energias além do petróleo, e a criação de condições para um desenvolvimento sustentável.

Na mesma linha, David Rothenberg, doutor em Filosofia, músico e professor do New Jersey Institute of Technology, denuncia a difícil situação das baleias jubarte no mundo e se aproxima para estabelecer diálogos com elas por meio da música. Seu álbum sobre as baleias jubarte é o registro da natureza mais vendido da história. Annette Bhagwati, doutora em Antropologia Social e especialista em África e Ásia, é internacionalmente reconhecida por seu trabalho sobre as culturas e suas transformações, por meios diferentes aos tradicionais. Problemas complexos como migração, discriminação racial e violência de gênero são observados por Bhagwati, por meio de perspectivas alternativas como oralidade e práticas estéticas. E por fim, Ricardo Dal Farra, um líder mundial em criação e pensamento ambiental, propõe um Equilíbrio-Desequilíbrio, um cenário que denuncia, ano após ano, as crises causadas pelas mudanças climáticas, por meio das interseções entre arte, natureza, ciência, tecnologia e sociedade.

I.J.: Como as HD se relacionam com o humanismo atualmente, ou seja, como se constituem em um instrumento de transmissão de valores humanísticos tradicionais às novas gerações e de promoção de novos valores necessários ao mundo de hoje?

F.L.: Gostaria de dar uma resposta curta a esta última questão e começar com uma crítica à concepção tradicional de “humanismo”. Do nosso ponto de vista, as Humanidades Digitais querem transcender esse pensamento capitalista e patriarcal homogeneizante centrado no antropocentrismo, que causou muitos dos desastres naturais de hoje.

Posturas como as de Stefano Mancuso, Rosie Braidotti, Arturo Escobar e tantos outros, nos convidam a compreender a complexidade das relações entre o humano e o não humano, entre o planetário e o natural, priorizando outras formas de ver de maneira mais igualitária, menos

hierárquica e mais comunal. As Humanidades Digitais nos permitem criar novas narrativas sobre o mundo que habitamos e nos apresentam perspectivas críticas sobre o papel do ser humano no mundo. Design, arte, ciência e tecnologia, integrados, nos levam a buscar alternativas que permitam o fluxo de forças vitais para as diversas interconexões, hoje fundamentais para a sobrevivência da habitabilidade planetária.

Para pensar além do antropocentrismo, o Festival Internacional da Imagem, que realizaremos na Colômbia de 24 a 28 de maio de 2021, propõe como tema central a investigação do conceito de Interespécies, não apenas como forma de compreensão da comunicação ou experimentação científica entre várias espécies, mas principalmente como forma de buscar relações vitais múltiplas e emergentes, a partir de novas formas de interação com os outros, humanos e não humanos. A capacidade geradora no continuum natureza-cultura, como gesto de afirmação das relações de proximidade espaço-temporal, suscita outras formas de expressão e ação próprias dessas interações com os outros, onde o poder da memória e da continuidade no tempo, em termos de sustentabilidade, possibilite um mundo habitável de forma digna para as gerações futuras.

I.J.: *Foi uma honra, Prof. Londoño. Muito obrigada!*

Referências

LONDOÑO, Felipe. *Expresión visual en las ciudades del Bahareque*. Colombia: Ed. Universidad de Caldas, 1994.

_____. Honrar la palabra, un ejercicio de coherencia en tiempos difíciles. *Site Universidad de Caldas*, 2018. Disponível em: <ucaldas.edu.co/portal/rector-u-de-caldas-honrar-la-palabra-un-ejercicio-de-coherencia-en-tiempos-dificiles/>. Acesso em 31 ago. 2020.

_____. *Interfaces de las comunidades virtuales*. Colombia: Ed. Universidad de Caldas, 2004.

_____. *¿Qué es diseño hoy? 1er Encuentro Nacional de Investigación en Diseño*. Colombia: Ed. Universidad Icesi Departamento de Diseño, 2004.

LONDOÑO, Felipe; CASTANEDA, Walter; RENDON, Guillermo; GOMEZ, Adriana. *Patrones de Color: Interpretación visual de los valores cromático regionales en Caldas*. Colombia: Centro Editorial Universidad de Caldas, 2006.

LONDOÑO, Felipe; GÓMEZ ALZATE, Adriana. *Paisajes y nuevos territorios (en red): cartografías e interacciones en entornos visuales y virtuales*. Barcelona: Anthropos, 2011.

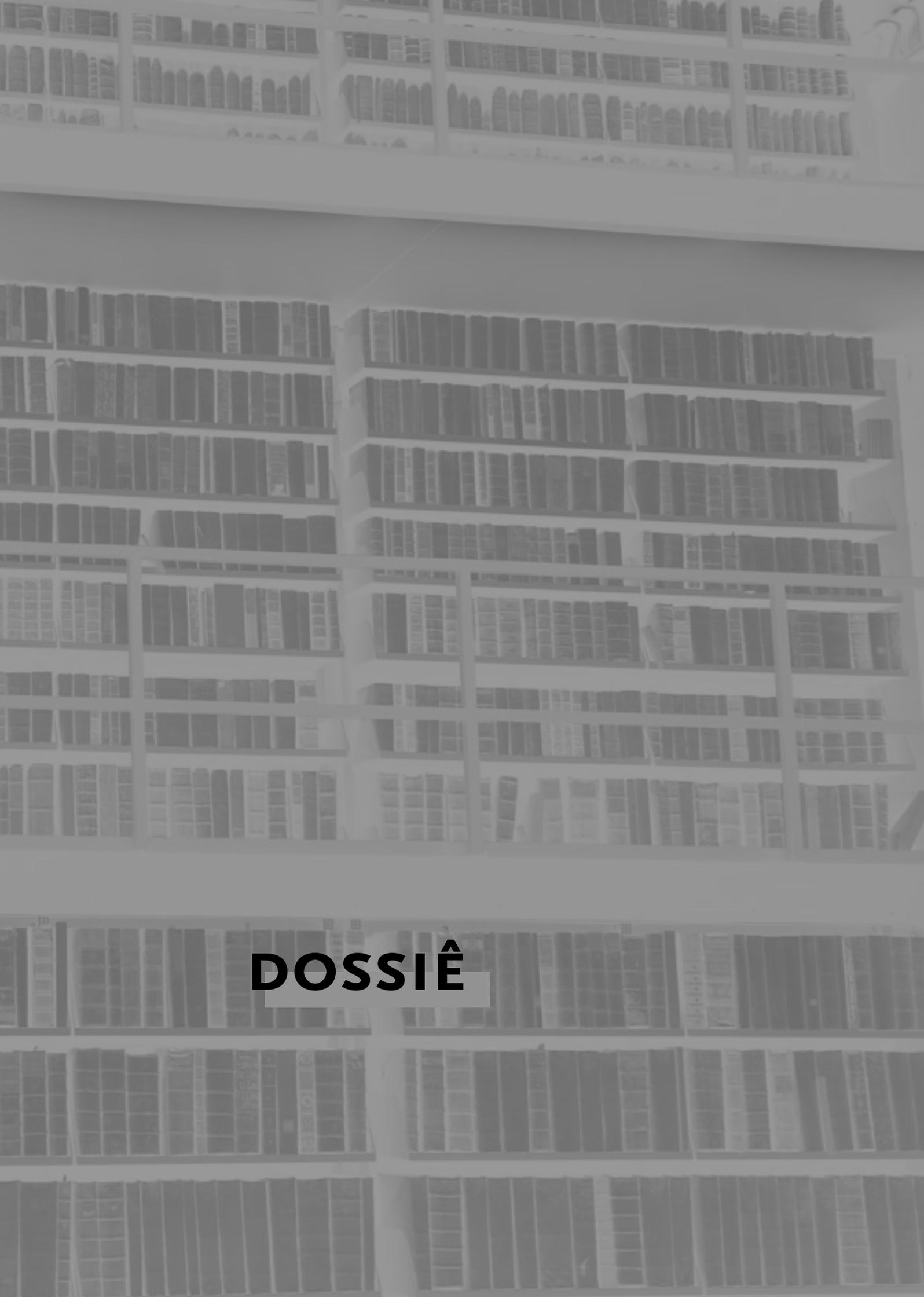
LONDOÑO, Felipe; MEJIA, German; CUESTA, Carlos; SANCHEZ, Yolima. *Videojuegos, diseño y ciudadanía*. Colombia: Ed. Universidad de Caldas, 2012.

LONDOÑO, Felipe; VALENCIA, Mario. *Metodología para la creación de proyectos interactivos*. Colombia: Ed. Universidad de Caldas, 2004.

_____. *Diseño digital*. Colombia: Ed. Universidad de Caldas, 2006 e 2009.

ROCHA, Cleomar; LONDOÑO, Felipe; VENTURELLI, Suzete. *Dimensões: arte_ design_ tecnologia*. Goiânia: CEGRAF/UFG, 2020.

WALLERSTEIN, Immanuel. *Abrir las Ciencias Sociales*. México: Siglo XXI Editores, 2006.



DOSSIÊ

Bases conceituais e tecnológicas para a modelagem de megadados em Humanidades Digitais:

um estudo de caso em *corpora* textuais de história da ciência

Ana Maria Alfonso-Goldfarb¹

José Luiz Goldfarb²

Márcia Helena Mendes Ferraz³

Odécio Souza⁴

Resumo: O intenso trabalho de digitalização realizado desde os anos 90 disponibilizou vastíssimos *corpora* textuais, inclusive para acesso remoto. Mais recentemente, novos recursos informáticos vêm fornecendo possibilidades inéditas de estudos teóricos, contemplando, também, a modelagem de megadados para organização e classificação dos conhecimentos. O presente estudo de caso se inscreve num marco maior de cooperação internacional, destinado a tornar mais efetiva a pesquisa e o trabalho documental em história da ciência. Especificamente, seu objetivo é elaborar ferramentas que permitam a localização e reconhecimento de conceitos comuns a grupos de textos pertencentes a grandes bases de dados em e para a história da ciência, assim como a sua mudança em função do tempo, tendo em vista sua indexação e classificação.

Palavras-chave: História da ciência. Humanidades Digitais. Organização e classificação do conhecimento. *Corpora* textuais. Modelagem e mineração de dados. Linguística computacional.

1 cv Lattes: lattes.cnpq.br/7427854657719431.

2 cv Lattes: lattes.cnpq.br/1023793876897710. E-mail: jlgoldfarb@dialdata.com.br.

3 cv Lattes: lattes.cnpq.br/5677877981039661.

4 cv Lattes: lattes.cnpq.br/4189566610252580. E-mail: odeciosouza@gmail.com.

Conceptual and technological grounds for big data modelling in digital humanities: textual corpora in history of science as case study

Abstract: Extensive digitization projects conducted since the 1990s made large textual corpora easily available, even to remote access. More recently new computing tools began to be developed that afford novel approaches to theoretical studies and big data modeling likely to allow for more accurate standards for the organization and classification of knowledge. The present case study is integrated within a larger international collaboration aiming at making work with documents relevant for history of science more effective. More in particular, the aim of the present case study is to develop tools to locate and recognize concepts common to texts included in large databases specific for history of science research, as well as their change over time, for the purpose of indexing and classification.

Keywords: History of science. Digital humanities. Organization and classification of knowledge. Textual corpora. Data mining and modeling. Computational linguistics.

Qualificação do problema

Como se sabe, na última década e de modo crescente, a tecnologia digital vem alterando a forma de planejar e executar o trabalho feito por pesquisadores. De fato, conforme Berry (2012), cada vez mais, a atividade de pesquisa é mediada pela tecnologia digital, a tal ponto que o próprio conceito de “pesquisa” vem mudando e, por consequência, afetando as epistemologias e ontologias que subjazem aos programas de pesquisa.

No caso das Ciências Humanas e Sociais – de particular interesse para nosso caso –, originalmente, desenvolveu-se um campo conhecido como “computação nas humanidades” (*humanities computing*). Este, porém, consistia na aplicação, pura e simples, de técnicas de computação aos próprios materiais e objetos gerados ou utilizados pelas Humanidades, visando reproduzir digitalmente grandes projetos editoriais. Todavia, não demorou muito para que as possibilidades geradas pelas novas tecnologias ultrapassassem essa perspectiva (BERRY, 2012; CARACO, 2012; HAYLES, 2012; MCCARTY, 2005; SVENSSON, 2009, 2010), dando início ao que hoje se conhece como “Humanidades Digitais” (HD). Como esperado, tal mudança implicou uma profunda transformação conceitual, resultando num novo campo do saber. Vale lembrar que a escolha do termo HD aponta muito mais para similaridades entre metodologias do que entre objetos, textos ou simples tecnologias (KIRSCHENBAUM, 2010). Assim, o termo é geralmente utilizado como uma espécie de guarda-chuva que abriga “um vasto conjunto de práticas para criar, aplicar, interpretar, interrogar e hackear tecnologias de informação, tanto as mais antigas quanto as mais novas” (PRESNER, 2010). Caberia, portanto, perguntar se, de fato, estamos tratando com uma única, ou com várias linhagens de HD (FITZPATRICK, 2011).

Nesse sentido, desde o início e conforme já indicado, foram percebidas diferenças na aplicabilidade das tecnologias digitais aos diversos campos do saber. Nada semelhante ao sucesso quase imediato e espetacular nas ciências exatas e naturais se deu entre as ciências humanas e sociais. Muito embora, no caso das pesquisas em ciências sociais a assimilação tenha sido mais fácil e breve do que entre as ciências humanas,

propriamente ditas, e de forma especial nas pesquisas em História (BAUR, 2009; WELDON, 2015; Aronova *et al.*, 2017). De acordo com a socióloga N. Baur (2009), os motivos para essas dificuldades seriam vários: 1) diferente dos sociólogos, os historiadores tradicionalmente têm uma percepção aguda de que a interpretação de seus dados depende, em boa parte, da perspectiva de análise tomada, bem como das especificidades próprias a diferentes períodos históricos e localizações geográficas. Ou, em outras palavras, o trabalho com dados históricos depende de algumas regras áureas para evitar interpretações normativas e enviesadas; 2) enquanto os sociólogos focam sua atenção em dados primários, resultantes, em geral, da pesquisa de campo, bem como em sua análise secundária (de maneira semelhante às ciências naturais), os historiadores buscam seus dados em bases documentais, com frequência resultantes de intrincados processos históricos e, portanto, difíceis de extrair e analisar; 3) enquanto os historiadores colocam grande atenção em verificar todo e cada um de seus dados, os sociólogos visam à obtenção de amostragens de dados que sejam abrangentes.

Em síntese, seria possível dizer que, diferentemente da abordagem de análise histórica, a sociológica tende a deixar de lado fatores individuais e contingentes, passíveis de produzir dados incomuns ou fora de esquadro que poderiam interferir no conjunto mais geral e modular pretendido (WELDON, 2015; ALFONSO-GOLDFARB *et al.*, 2018). Todavia, nos últimos anos, as pesquisas históricas têm buscado um nicho adequado e próprio nas HD. No que diz respeito ao presente estudo de caso, seria importante indicar que, já há algum tempo, pesquisadores em História e Filosofia da Ciência (HFC), assim como em estudos em ciência e tecnologia (ECT) têm buscado conceitualizar as várias ciências e suas histórias como um sistema complexo que inclui uma diversidade de fatores. Tais fatores abrangem desde redes sociais e institucionais, padrões de financiamento, desenvolvimentos técnicos e tecnológicos, até as variadas mudanças nos pressupostos epistemológicos, passando por sua lógica intrínseca e outras questões próprias ao conhecimento científico, tal como refletidas nos *corpora* textuais.⁵ Até o momento, é possível dizer que vários desses fatores têm se mostrado passíveis de modelagem e análise digital.

5 *Corpus* textual (plural, *corpora*), na linguística, é um conjunto amplo e estruturado de textos, utilizado para a realização de análises estatísticas e testar hipóteses. Assim, essa expressão passou a ser utilizada para nomear o substrato de trabalho nas Humanidades Digitais; grosso modo corresponde ao que nas disciplinas históricas é denominado *corpus* documental. Para maior detalhe, vide Wynne (2005).

Contudo, esse encontro, entre a nova tecnologia digital e as Humanidades, não ocorreu sem uma certa tensão. O motivo fundamental é que nas HD, em geral, e nas áreas históricas, em particular, a geração de conjuntos de dados (*datasets*) não antecede nem é independente do próprio trabalho de pesquisa (SCHREIBMAN *et al.*, 2008, p. xxv). Assim, van Zundert (2012, p. 173-174) alerta que as ferramentas informáticas “também devem assegurar que as heurísticas e hermenêuticas já existentes sejam apropriadamente traduzidas nos seus equivalentes digitais, especialmente nas áreas nas quais a heterogeneidade dos dados e abordagens multifacetadas não são consideradas ruído a ser reduzido, mas propriedades essenciais da área”.

Não por acaso, a maioria dos projetos ainda mantêm o foco em aspectos mais pertinentes às ciências sociais, tais como redes sociais, financeiras, tecnológicas e institucionais. Muito embora, o escopo de aplicação de tecnologias digitais à pesquisa em história da ciência já consegue abranger diversas possibilidades, incluindo a análise estatística e semântica de seus *corpora* textuais, de sorte que a criação de acervos digitais organizados nesse campo – tanto com base na vida e na obra de cientistas individuais, como Isaac Newton (WALSH & HOOPER, 2012; ILIFFE, 2004; PASTORINO *et al.*, 2008), Henri Poincaré (*Poincaré Project*) e Charles Darwin (*Darwin Correspondence Project*, *Darwin Manuscript Project*), quanto de instituições, como *History of the Marine Biological Laboratory* ou *History of the Max Planck Society* – vêm estabelecendo uma linha bem definida de estudos. A organização e catalogação de tais acervos, especialmente de correspondências e atas de sessões institucionais, já teve como efeito colateral a produção de uma grande riqueza de metadados relacionais, que permitiram descrever redes de interação entre atores, sua produção coligada e instituições de referência mais significativas. Outros bons exemplos nessa direção são os projetos *Six Degrees of Francis Bacon* e *Registres de l'Académie*, cujo tipo de *dataset* gerado já se assemelha ao obtido pelas Ciências Sociais.

Esse intenso trabalho de digitalização de documentos *em e para* a história da ciência, realizado desde os anos 90, vem fornecendo também novas possibilidades de estudos teóricos sobre a organização e classificação dos conhecimentos. Hoje considerados de grande relevância para o desenvolvimento de campos em formação como as HD, esses estudos são o ponto focal do presente estudo de caso, conforme será visto a seguir.

Foco de nossas pesquisas em Humanidades Digitais (HD), História e Filosofia da Ciência (HFC) e Estudos de Ciência e Tecnologia (ECT)

O foco de nossas pesquisas tem como base trabalhos realizados ou em desenvolvimento no CESIMA (Centro Simão Mathias de Estudos em História da Ciência, PUC-SP), centro de pesquisa fundado em 1994. Apenas para recordar, de forma breve, nosso centro de pesquisa foi criado, explicitamente, dentro de uma perspectiva digital, logo no início dos questionamentos sobre a aplicabilidade dessa nova vertente aos complexos campos das Humanidades. O objetivo imediato era disponibilizar, em suporte digital, um *corpus* documental significativo para pesquisas em história da ciência e campos relacionados. Desde o início, a biblioteca digital do CESIMA esteve aberta, não só a membros de nosso centro, mas a outros estudiosos brasileiros ou latino-americanos que requeressem seu uso para fins de pesquisa. Vale lembrar que, antes da possibilidade recente de acesso pela rede mundial, a consulta aos documentos originais com frequência demandava o deslocamento dos pesquisadores até outros centros, muitas vezes distantes, como os europeus, norte-americanos ou mezzorientais (GOLDFARB & SOUZA, 2018). Ou ainda, aqueles centros que, tanto no país, quanto no exterior, não tinham seus acervos microfilmados ou digitalizados, trabalho este iniciado com sucesso por membros de nossa equipe. Assim, a primeira fase do trabalho consistiu na digitalização de um vasto *corpus* textual, que, na atualidade, compreende mais de 30.000 títulos. A segunda fase, já em vias de conclusão, consiste na disponibilização completa do acervo para acesso remoto.

Contudo, ao organizar a base digital de dados, nos deparamos com uma série de questões, relacionadas à classificação e à catalogação dos documentos, que dificultavam bastante a sua localização pelos usuários. Questões que, na verdade, não são de fácil solução, pois no caso da história da ciência têm origem em sua própria natureza, estabelecida através das interfaces entre um variado leque de áreas (ALFONSO-GOLDFARB, 2008; ALFONSO-GOLDFARB *et al.*, 2013; GARBER, 2002, 2009; RHEINBERGER, 2007, 2010; WELDON, 2009; WITHROW, 1964, 1976). Por sua vez, tal singularidade pode levar ao fenômeno de *overtagging*, em buscas digitais, em que o volume e emaranhado de termos obtidos é tal que não se produzem resultados significativos (HANRAHAN & SCHNÖPF, 2013). Com a finalidade de enfrentar essa e outras questões, referentes à organização e classificação do conhecimento, o CESIMA deu início a uma série de pesquisas e atividades envolvendo parcerias nacionais e internacionais. Dentre estas, vale destacar as mantidas há muitos anos com o *Grupo de Pesquisa em Tec-*

nologia Aplicada à Educação (GTech.Edu/UFRGS), produtor do minerador de conceitos brasileiro *Sobek*, o *Digital History and Philosophy of Science Consortium* (dHPS), o *Committee of Bibliography and Documentation, Division of History of Science and Technology, International Union of History and Philosophy of Science* (CBD/DHST/IUHPS) e a *Isis Current Bibliography* (ICB) o principal indexador de nosso campo de estudos.

Nesse contexto, inicialmente, foi desenvolvido, já em 1994, um projeto piloto para criação de biblioteca digital que, então inédita no meio acadêmico nacional, teve apoio da FAPESP. O fortalecimento e ampliação dessa iniciativa contou por quase duas décadas (1999-2018) com três grandes Projetos Temáticos da FAPESP. De forma mais específica e, desta vez, com apoio do CNPq, entre os anos de 2010-2012, foi desenvolvido o projeto “Novas perspectivas de classificação e abordagem em história da ciência: aspectos teórico-metodológicos e técnicos para elaboração de instrumentos adequados de busca”. No decorrer do mesmo, foram realizadas revisões meticulosas e, ao mesmo tempo, abrangentes das assim chamadas “árvores do conhecimento”, ou sistemas de hierarquização das ciências. Ao mesmo tempo, foram também realizados encontros programados, onde se deram amplas discussões com destacados especialistas em classificação, bibliografia, biblioteconomia, arquivística, além de especialistas em ciências da informação e em tecnologia da informação (TI). Os resultados, comunicados em diversos encontros internacionais e publicações especializadas, nos levaram a propor o desenvolvimento, para uma base digital em história da ciência, da chamada classificação *Colon*. Tal classificação foi elaborada nos anos 30, do século xx, pelo matemático indiano S.R. Ranganathan e aplicada com grande sucesso nas Ciências Exatas e Naturais, a partir da década de 1950, pelo *Classification Research Group* (CRG) do Reino Unido (FRICKÉ, 2012, p. 212 *et seq.*; SPITERI, 1998).

O sistema de Ranganathan apresenta notáveis vantagens: 1) decompõe (faz a análise) a informação em unidades básicas (“facetar”, correspondendo a: tempo, espaço, materialidade, energia/movimento e personalidade/características individuais) que, na sequência, podem ser recombinadas (faz a síntese), dando origem a categorias gerais, incluindo as multi-, inter- e transdisciplinares; 2) a abordagem “de baixo para cima” (ou seja, de elementos inerentes ao próprio documento às categorias gerais às quais pertence) que acompanha a atuação real do usuário, especialmente no mundo digital, no qual a sequenciação temporal (*strings*) substituiu a categorização espacial (prateleiras tangíveis em bibliotecas); 3) as facetar permitem representar conceitos, em vez de termos (RANGANATHAN, 2012).

Todavia, é preciso lembrar que, pelo menos, desde a década de 1960 entende-se que a história da ciência lida com conceitos e suas ressignificações, em vez de termos (CANGUILHEM, 1977, p. 11-27). Assim, as tentativas iniciais logo detectaram dois obstáculos de difícil solução: 1) um dos passos na programação exige, necessariamente, decisão humana (escolha de palavras-chave mais significativas ou que melhor descrevam um grupo de documentos) e, portanto, não poderia ser automatizado; 2) levando em conta, como indicado há pouco, que a história da ciência lida com conceitos e que estes mudam no decorrer do tempo, seria essencial incluir esse componente de variação nos demais itens da classificação. Portanto, seria necessário elaborar uma modelização do tempo, algo que até o momento é de difícil realização, como demonstrado em publicações recentes de, por exemplo, Weldon (2015); Alfonso-Goldfarb *et al.* (2018); Gibson *et al.* (2019).

Conforme indicado antes, porém, existem outras possibilidades que implicam na adequação de modelos e métodos desenvolvidos em outros campos informáticos às questões próprias da história da ciência. Uma de tais estratégias é a modelagem de tópicos, a saber, um conjunto de métodos matemáticos e estatísticos para se inferir conceitos ou tópicos a partir de grandes coleções de textos digitalizados. Tais métodos foram desenvolvidos dentro da área de aprendizagem automática (*machine learning*) com o objetivo básico de otimizar ferramentas de busca *online*. A modelagem de tópicos foi amplamente adotada nas HD, geralmente para explorar o conteúdo temático de grandes *corpora* históricos ou literários. Em HFC/ECT, a modelagem de tópicos pode ser utilizada para mapear as mudanças na representação de tópicos e conceitos através do tempo (HALL *et al.*, 2008). Como exemplos dessas ferramentas modeladoras, temos a *Análise Semântica Latente* (LSA – *Latent Semantic Analysis*) e a *Alocação Latente Dirichlet* (LDA – *Latent Dirichlet Allocation*), aplicadas para mapear, respectivamente, a cronologia dos trabalhos alquímicos/químicos por Newton (WALSH & HOPPER, 2011) e as mudanças dos hábitos de leitura de Darwin (MURDOCK *et al.*, 2015). Em ambos os casos, trata-se de uma análise quantitativa que expressa relações latentes entre textos e conceitos.

De maneira semelhante, o campo da linguística computacional representa uma fonte promissora para a abordagem digital da documentação em história da ciência. A linguística computacional utiliza a aprendizagem automática de maneira estatística para estudar a estrutura e a evolução da linguagem, permitindo a detecção e comparação de estruturas epistêmicas e padrões de linguagem numa determinada comunidade científica. Um caso exemplar é representado pelo trabalho desenvolvido por Pumfrey & Ashcroft (2015) sobre a conceptualização dos trabalhos de laboratório na primeira modernidade.

Assim, para melhor indicar as condições existentes para o desenvolvimento e estabelecimento dessas pesquisas, junto ao CESIMA, na sequência são oferecidos dois itens especificando o trabalho de implementação da biblioteca *CESIMA Digital* e seus vínculos com outros centros de estudos.

a. Biblioteca *CESIMA Digital*

Originalmente dedicada a digitalizar obras antigas ou raras sobre as ciências, em diferentes épocas e suportes (manuscritos, livros, cartas, diários, atas de reuniões, cadernos de laboratório e outros mais em microfílm, microformas, papel etc.), a base de dados documental do CESIMA passou por uma série de modificações, através da adaptação de equipamentos e softwares diversos, até tornar-se a atual Biblioteca *CESIMA Digital*.

No que tange ao acervo, este foi obtido a partir de autorizações ou aquisições nos mais diversos centros de documentação, bibliotecas, arquivos, nacionais e internacionais. A padronização dessa documentação digitalizada ou já digital foi feita em Adobe/PDF, por ser este considerado um padrão mundialmente utilizado. Além disso, todos os arquivos com origem em material impresso foram submetidos a aplicativo de reconhecimento de caracteres (OCR – *Optical Character Recognition*), de maneira a facilitar as buscas e uso de mineradores. Todavia, a Biblioteca *CESIMA Digital* também engloba um número considerável de arquivos com origem em manuscritos que ainda depende de futuros desenvolvimentos de aplicativos especiais para reconhecimento de seus irregulares e variadíssimos tipos de caracteres.

Atualmente, o acervo completo, passado por OCR ou não, encontra-se no servidor do CESIMA, instalado e configurado no *backbone* na *Divisão de Tecnologia da Informação* (DTI) da PUC-SP. Foi selecionado o software *Dspace* de gerenciamento e preservação de documentos online, sendo este compatível com um sistema internacional de suporte a bibliotecas e indexação de seus conteúdos. Para sua utilização, dispõe-se de duas unidades de armazenamento (*Storage*) conectadas que se encontram devidamente instaladas e ligadas ao servidor do CESIMA com autorização de acesso pela DTI. Além dessa configuração inicial, foi escolhido o banco de dados gratuito *PostgreSQL*, compatível e adequado ao *Dspace*. Detalhes teóricos e técnicos, mais específicos, podem ser obtidos em Souza (2019), cuja tese de Doutorado teve como foco o processo de formação e disponibilização da Biblioteca *CESIMA Digital*.

Aberta universalmente ao público, em julho de 2018, a Biblioteca *CESIMA Digital*, oferece mais de 1/3 de seu acervo para acesso remoto, enquanto o restante passa por um criterioso, mas acelerado, processo de verificação e catalogação antes de seu *upload* que, se espera, seja concluído brevemente. Embora o catálogo desse grande conjunto de obras seja aberto, para acessá-las *online* ou fazer seu *download* é solicitado aos usuários um cadastramento, a partir do qual recebem uma senha de acesso, após aprovação do comitê científico do *CESIMA*. Desta forma, tornou-se possível garantir o devido respeito aos direitos autorais e uso científico dos materiais acessados, conforme previamente solicitado pelos centros de documentação em que tiveram origem e a norma internacionalmente aceita. De igual maneira, tal cadastramento tem facilitado saber quem são os usuários, de onde provêm e, de certa forma, suas áreas de interesse. A relevância de tal iniciativa pode ser avaliada por seu alcance. Com menos de dois anos de operação, ainda sem o catálogo completo *online* e sem grande divulgação, a Biblioteca *CESIMA Digital* já conta com dezenas de usuários cadastrados e centenas de consultas, somando 1946 acessos, desde 97 diferentes territórios, em diferentes partes do mundo. No mapa abaixo, as áreas em azul indicam os territórios originários das consultas, enquanto a tabela subsequente oferece sua localização específica e seus números, até os primeiros meses de 2020.

Todavia, lembrando que o *CESIMA Digital* não se trata apenas de uma biblioteca online para consulta, como tantas outras, mas de uma coleção muito especial reunida por pesquisadores – através de suas pesquisas individuais e conjuntas – e oferecida a outros estudiosos de HFC/ECT, as questões teóricas e práticas ali encetadas têm sido de grande interesse para as HD e campos afins. Nesse sentido, o próximo item é dedicado,



Figura 1. No mapa, as áreas em azul indicam os territórios originários das consultas à Biblioteca Digital do *CESIMA* (Fonte: AUTORES, 2020)

Acessos por País (01jul2018 a 19mar2020) total 1946					
Brazil	829	Russia	8	Pakistan	3
United States	192	Austria	7	Romania	3
Argentina	178	Chile	7	Slovenia	3
United Kingdom	77	Colombia	7	Thailand	3
Portugal	69	Greece	7	Georgia	2
Spain	48	Ireland	7	Hong Kong	2
Italy	43	Sweden	7	Iraq	2
France	33	Czechia	6	Kenya	2
Mexico	32	Egypt	6	Lebanon	2
Germany	29	Hungary	6	Sri Lanka	2
China	26	South Africa	6	Mozambique	2
Canada	25	Ecuador	5	Norway	2
Netherlands	24	Finland	5	Slovakia	2
India	20	Morocco	5	El Salvador	2
Australia	19	Poland	5	Ukraine	2
Japan	17	Taiwan	5	Venezuela	2
(not set)	15	Denmark	4	United Arab Emirates	1
Switzerland	12	New Zealand	4	Angola	1
Israel	12	Peru	4	Bosnia & Herzegovina	1
Philippines	12	Cameroon	3	Bulgaria	1
Turkey	11	Algeria	3	Benin	1
Indonesia	10	Ghana	3	Belize	1
South Korea	10	Iran	3	Côte d'Ivoire	1
Belgium	8	Malaysia	3	Costa Rica	1
				Zimbabwe	1

Figura 2. A tabela mostra os acessos por país das consultas e seus números até os primeiros meses de 2020. (Fonte: AUTORES, 2020)

especialmente, a oferecer um pouco mais de detalhes sobre duas parceiras, dentre aquelas mantidas pelo CESIMA, com foco particular em sua biblioteca digital.

b. Parcerias focadas na Biblioteca CESIMA Digital

Entre os parceiros mais relevantes para o *CESIMA Digital*, indicados anteriormente – CBD, ICB, GTech.Edu, dHPS –, os dois últimos merecem especial atenção pelo trabalho conjunto que vem sendo desenvolvido no que concerne as HD e campos coligados. O mais antigo dentre ambos, com cuja equipe já desenvolve pesquisas há anos, é o GTech.Edu, estabelecido na UFRGS e reconhecido nacional e internacionalmente.

Liderada pelo Prof. Eliseo Reategui, essa equipe desenvolveu o SOBEK, ferramenta de acesso aberto que fornece a possibilidade – não só a pesquisadores da UFRGS, mas a outros espalhados pelo planeta – de compor, interpretar e avaliar textos. Segundo indicado na própria página de tal ferramenta, seus fundamentos técnicos e teóricos foram desenvolvidos a partir dos estudos de Schencker (2003), cuja tese de Doutorado fornece os princípios estatísticos a partir dos quais os algoritmos são habilitados para “minerar”, ou seja, identificar os termos mais significativos de um texto e, seguidamente, relacioná-los em “árvore”. Tal processo, utilizado e ampliado pelo SOBEK, transformaria simples termos em *quasi*-conceitos, uma vez que passa a fornecer sentido a estes, através de seus vínculos hierárquicos e relacionais com o todo do texto, possibilitando, assim, a sua interpretação.

Uma nova versão do SOBEK para *desktop*, elaborada em conjunto com o Prof. Reategui e sua equipe, possibilitou a obtenção de um bom número desses *quasi*-conceitos em uma seleção de obras pertencentes ao acervo do *CESIMA Digital*. Com isso, encontra-se em desenvolvimento um trabalho para, a partir desses *quasi*-conceitos, formular computacionalmente o que especialistas em documentação chamam de descritores, ou seja, descrições breves de cada documento associadas às informações convencionais para a sua catalogação (autor, título, ano etc.). Penosos e demorados, pois feitos individualmente por especialistas e/ou documentalistas conhecedores de cada obra em questão, os descritores quase sempre aparecem em catálogos virtuais apenas como cópias dos já existentes em meio físico. Dessa forma, a nova versão do SOBEK, aplicada ao *CESIMA Digital*, deverá oferecer, em futuro não muito distante, possibilidades de busca que vão além das meras informações dos catálogos convencionais. Possibilidades essas que poderão se ampliar imensamente uma vez cruzadas, por tal sistema, às informações contidas em diferentes obras e, assim, verificar seus conjuntos conceituais em comum. Algo que, conforme já mencionado, interessa de forma particular aos estudos em história da ciência, sempre focados em identificar a mudança conceitual de um mesmo termo, em diferentes épocas. Em uma perspectiva mais próxima e um pouco menos ambiciosa, já se encontra em desenvolvimento um trabalho para, a partir dos descritores, acompanhar – com a devida autorização do usuário – quem consultou ou está consultando um determinado documento, quais outros documentos semelhantes consultou, e se deixou alguma observação sobre os mesmos (SOUZA, 2019).

Uma segunda parceria que vem se desenvolvendo com força, embora há menos tempo, está vinculada ao dHPS, por constituir-se num marco referencial para os estudos em HD. Marco este cujo espectro internacional e mais amplo vem, inclusive, beneficiando as parcerias já existentes e gerando outras novas para o *CESIMA Digital*.

Apenas para resumir algo de sua história, o dHPS foi criado em 2011, a partir de projetos em documentação exclusivamente para a história da ciência, financiados pela *Nacional Science Foundation (NSF)* dos EUA. Não demoraria muito para que esse consórcio conseguisse reunir tanto historiadores e filósofos da ciência, quanto especialistas em TI, Ciências da Computação, Biblioteconomia e outras ciências da informação. O propósito original desse colegiado multi e interdisciplinar era desenvolver novas formas de integrar as abordagens tradicionais de pesquisa, em HFS, com ferramentas computacionais e recursos digitais. Atualmente, a missão proposta pelo dHPS é mais abrangente e ambiciosa, conforme

expresso em suas próprias palavras: “desenvolver, dar suporte e promover projetos digitais em HFC, incluindo edições, publicações e as ferramentas de pesquisa necessárias para tanto. Na medida do possível e reconhecendo os desafios e limites o Consórcio está compromissado com produtos de código e acesso abertos. Comprometemo-nos a desenvolver infraestrutura sustentável para ambos, projetos e produtos” (Disponível em: digitallhps.org. Acesso em: 25 jul. 2020).

Estimulado pela própria NSF, entre 2016 e 2017, o dHPS apresentou e foi contemplado, por essa mesma agência, com um projeto para coordenar pesquisas em rede, o chamado *Research Coordination Network* (RCN). Fazem parte desse projeto, que se propõe a fundamentar boa parte dos trabalhos da dHPS, além de membros de diversas universidades dos EUA, também representantes das seguintes instituições: *Chemical Heritage Foundation/EUA*, *American Science Museum/EUA*, *Stevens Institute of Technology/EUA*, *Humboldt-Universität zu Berlin/Alemanha*, *Max Planck Institute/Alemanha*, *University of Melbourne/Austrália*, *University of Cambridge/Reino Unido*, *Université de Nantes/França*. O CESIMA/PUC-SP teve o prazer de também ser convidado a participar desse seleto grupo, colaborando com a elaboração do projeto desde o início e até agora com o seu desenvolvimento.

A partir desse projeto, um dos desafios específicos do dHPS é desenvolver *standards* e soluções tecnológicas para mapear e integrar as diferentes autoridades e ontologias em uso. Primeiramente, espera-se que a boa consecução desse processo deva alcançar o tão sonhado compartilhamento de dados entre projetos. Mas, além disso, a pesquisa necessária para tanto, também deve produzir novas bases conceituais e epistemológicas, no que diz respeito ao manejo e à modelagem de megadados (*big data*), para o trabalho documental em HFC/ECT (GIBSON *et al.*, 2019). Nesse contexto, o *CESIMA Digital*, ao não consistir em meras coleções de HFC/ECT, mas ter sido planejado, desde o início, como uma base de dados em que o texto e o contexto das obras devem ser vistos de maneira interligada, foi selecionado como um estudo de caso preferencial para as pesquisas do RCN. De igual forma, foi considerado um ambiente experimental ideal para testar e/ou desenvolver instrumentos de busca que visem detectar conceitos e suas transformações, ao longo do tempo, em textos dedicados às ciências de diferentes épocas.

Considerações finais

A pesquisa em Humanidades, progressivamente inserida na era digital, vem se transformando, ao aderir, não só aos processos tecnológicos, mas teórico-metodológicos produzidos nesse novo contexto. Congregadas cada vez mais no campo recente das Humanidades Digitais (HD), pesquisas históricas, como aquelas em história da ciência e áreas afins, podem agora (ou poderão em breve) se beneficiar das novas ontologias e técnicas informáticas e produzir análises estatísticas e semânticas inéditas em seus *corpora* textuais.

De sorte que, ao formar parte de um marco internacional e mais amplo, em que esses estudos têm acontecido em ritmo célere, o *CESI-MA Digital* busca constituir uma via de mão dupla, através da qual possa fluir tanto o que for produzido de mais inédito e/ou recente no exterior, quanto as colaborações e/ou formulações originais brasileiras. Assim, espera-se que o desenvolvimento dos projetos dos quais participa coopere na elaboração de novas bases conceituais e tecnológicas referentes ao manejo e à modelagem de megadados para o trabalho documental das Humanidades. Todavia, considerando-se que o processo acima descrito tem como foco principal os estudos históricos, os resultados iniciais deverão contemplá-los e, de forma mais direta, aqueles relacionados à história da ciência e campos coligados. Espera-se ainda que esses desenvolvimentos promovam a vinculação e trabalho conjunto entre membros das equipes inter e multidisciplinares, nacionais e internacionais, que até agora não fazem parte de seu escopo atual de parcerias. Com isso, busca-se também a ampliação de tais redes de pesquisa, além da formação de novos pesquisadores e, não menos importante, um trabalho conjunto e potencializado para tornar efetivo o compartilhamento de dados entre projetos, assim como uma interface mais efetiva e sinérgica entre as pesquisas em Humanidades Digitais (HD) e aquelas em História e Filosofia da Ciência (HFC) e Estudos em Ciência e Tecnologia (ECT).

Referências

ALFONSO-GOLDFARB, A.M.; WAISSE, S.; FERRAZ, M.H.M. From shelves to cyberspace: the organization of knowledge and the complex identity of history of science. *Isis*, v. 104, n. 3, p. 551-560, 2013.

_____. New proposals for organization of knowledge and their role in the development of databases for history of science. *Circumscribere*, v. 21, p. 1-12, 2018.

- ALFONSO-GOLDFARB, A.M. Centenário Simão Mathias: Documentos, métodos e identidade da história da ciência. *Circumscribere*, v. 4, p. 5-9, 2008.
- ARONOVA, E.; VON OERTZEN, C.; SEPKOSKI, D. (org.). Data histories. *Osiris*, v. 32, n. 1, p. 1-354, 2017.
- BAUR, N. Problems of linking theory and data in historical sociology and longitudinal research. *Historical Social Research*, v. 34, n. 1, p. 7-21, 2009.
- BERRY, D.M. (org.) *Understanding digital humanities*. New York: Palgrave Macmillan, 2012.
- BERRY, D.M. The computational turn: thinking about the digital humanities. *Culture Machine*, v. 12, 2011.
- CANGUILHEM, G. *Ideologia e racionalidade nas ciências da vida*. Lisboa: Edições 70, 1977.
- CARACO, B. Les digital humanities et les bibliothèques. *Bulletin des bibliothèques de France*, n. 2, 2012.
- DEEGAN, M.; MCCARTY, W. (org.) *Collaborative research in the digital humanities*. Farnham: Ashgate, 2012.
- FERRAZ, M. H. M.; ALFONSO-GOLDFARB, A.M.; WAISSE, S. Science and History of Science: between Comte and Canguilhem. *Transversal*, v. 4, p. 108-117, 2018.
- FITZPATRICK, K. The humanities, done digitally. *The Chronicle of Higher Education*, 8/5/2011. Disponível em: <chronicle.com/article/The-Humanities-Done-Digitally/127382/>. Acesso em: 12 maio 2020.
- FRICKÉ, M. *Logic and the organization of information*. New York: Springer, 2012.
- MCCARTY, W. *Humanities computing*. New York: Palgrave, 2005.
- GARBER, D. Philosophia, historia and mathematica: shifting sands in the intellectual geography of the seventeenth century. *Studies in the History and Philosophy of Science*, v. 24, p. 1-7, 2009.
- GARBER, D. *Storia della scienza, v. 4: La rivoluzione scientifica*. Roma: Istituto della Enciclopedia Italiana, 2002.
- GARDINER, E.; MUSTO, R. *The digital humanities: a primer for students and scholars*. Cambridge: Cambridge University Press, 2015.
- GIBSON, A.; LAUBICHLER, M.D.; MAIENSCHIN, J. (org.). Dossiê: Computational History and Philosophy of Science. *Isis*, v. 110, n. 3, p. 497-566, 2019.

GOLDFARB, J.L.; SOUZA, O. From the Golem's Jewish Myth to IBM's responsive Watson: where are we going? *Circumscribere*, vol. 21, p. 118-122, 2018.

HANRAHAN, E.; SCHNÖPF, M. Scholarly digital editions: connecting archives and libraries. Conference *New Directions in Digital History of Science*. Berlin: Max Planck Institute of History of Science/Committee of Documentation and Bibliography, IUHPS, 2013.

HALL, D.; JURAFSKY, D.; MANNING, C.D. Studying the history of ideas using topic models. *Proceedings of the Conference on Empirical Methods in NATURAL Language Processing, EMNLP/2008*, Edimburgo, 25-27 outubro 2008.

HAYLES, K. *How we think: digital media and contemporary technogenesis*. Chicago: The University of Chicago Press, 2010.

HU, W.C.; KAABOUCH, N. *Big data management, technologies and applications*. Hershey: Information Science Reference, 2014.

HU, Y.; BOYD-GRABER, J.; SATINOFF, B.; SMITH, A. Interactive topic modeling. *Machine Learning*, v. 95, p. 423-469, 2014.

ILIFFE, R. Digitizing Isaac: the Newton Project and an electronic edition of Newton's papers. In: FORCE, J.E.; HUTTON, S. (org.) *Newton and Newtonianism: new studies*. Dordrecht: Springer, p. 23-38, 2004.

KIRSCHENBAUM, M. What is digital humanities and what's doing in English departments? *ADE Bulletin*, n. 150, p. 1-7, 2010.

MONTEMURRO, M.A.; ZANETTE, D.H. Keywords and co-occurrence patterns in the Voynich manuscript: An information-theoretic analysis. *Plos One* 21/6/2013 DOI: [10.1371/journal.pone.0066344](https://doi.org/10.1371/journal.pone.0066344)

MURDOCK, J.; ALLEN, C.; DEDEO, S. Exploration and exploitation of Victorian science in Darwin's reading notebooks. *Cornell University Library*, 10/12/2015 Disponível em: <[arXiv:1509.07175v2](https://arxiv.org/abs/1509.07175v2)>. [cs.CL]. Acesso em: 12 maio 2020.

PASTORINO, C.; LOPEZ, T.; WALSH, J.A. The digital Index Chemicus: toward a digital tool for studying Isaac Newton's Index Chemicus. *Body, Space & Technology*, v. 7, n. 2, 2008.

PRESNER, T. Digital humanities 2.0: a report on knowledge. *OpenStax CNX*. 8/6/2010. Disponível em: <cnx.org/contents/2742bb37-7c47-4bee-bb34-of35bda76of3@6>. Acesso em: 14 maio 2020.

PUMFRY, S.; ASHCROFT, P. Alchemy, chemistry or chymistry: an analysis of actors' categories in early modern England. *16th SHAC Postgraduate Workshop*. Oxford, 30 outubro 2015. Programme and Edited Abstracts, p. 5-6.

RANGANATHAN, S.R. *Colon Classification*. [reimp. sexta ed. 1960]. Bangalore: Ess Ess Publ., 2012.

RHEINBERGER, H.-J. *On historicizing epistemology: An essay*. Stanford (CA): Stanford University Press, 2010.

RHEINBERGER, H.-J. *Historische Epistemologie zur Einführung*. Hamburg: Junius, 2007.

SCHREIBMAN, S., SIEMENS, R.; UNSWORTH, J. (org.). *Companion to digital humanities*. Oxford: Blackwell, 2004.

SCHENCKER, A. *Graph-theoretic Techniques for Web Content Mining*. PhD Dissertation. Tampa: University of South Florida, 2003.

SOUZA, O. *Cesima Digital: uma ferramenta para a história da ciência*, 2019. Tese (Doutorado em História da Ciência). Pontifícia Universidade Católica de São Paulo, São Paulo.

SPITERI, L. A simplified model for facet analysis. *Canadian Journal of Information and Library Science*, v. 23, p. 1-30, 1998.

SVENSSON, P. Humanities computing as digital humanities. *Digital Humanities Quarterly*, v. 3, n. 3, 2009.

SVENSSON, P. The landscape of digital humanities. *Digital Humanities Quarterly*, v. 3, n. 3, 2009.

TERRAS, M; NYHAN, J.; VANHOUTTE, E. (org.) *Defining digital humanities: a reader*. Farnham: Ashgate, 2013.

UNSWORTH, J. What is humanities computing and what is not? In: TERRAS, M; NYHAN, J.; VANHOUTTE, E. (org.) *Defining digital humanities: a reader*. Farnham: Ashgate, p. 35-48, 2013.

WALSH, J.A.; HOOPER, W.E. The liberty of invention: alchemical discourse and Information Technology standardization". *Literary and Linguistic Computing*, v. 27, p. 55-79, 2012.

WELDON, S.P. Historians and their data. In: ALFONSO-GOLDFARB, A.M. et al. (org.) *Crossing oceans: Exchange of products, instruments and procedures in the history of chemistry and related sciences*. Campinas: CLE/UNICAMP, p. 299-322, 2015.

WELDON, S.P. The Isis bibliography from its origins to the present day: one hundred years of evolution of a classification system. *Circumscribere*, v. 6, p. 26-46, 2009.

WITHROW, M. A classification scheme for the history of science, technology and medicine. *Isis Cumulative Bibliography*, v. 3, p. 621-623, 1976.

WITHROW, M. Classification schemes for the history of science: a comparison. *Journal of Documentation*, v. 20, n. 3, p. 120-136, 1964.

WYNNE, M. (org.). *Developing linguistic corpora: a guide to good practice*. Oxford: Oxbow Books, 2005.

ZUNDERT, J. v. "If you build it, will we come? Large scale digital infrastructures as a dead end for digital humanities". *Historical Social Research*, v. 37, n. 3, p. 165-186, 2012.

Consórcios, projetos e ferramentas

BIBLIOTECA CESIMA Digital. Disponível em: <cesimadigital.pucsp.br>. Acesso em: 25 jul. 2020.

DARWIN CORRESPONDENCE PROJECT. Disponível em: <darwinproject.ac.uk>. Acesso em: 25 jul. 2020.

DARWIN MANUSCRIPT PROJECT. Disponível em: <amnh.org/our-research/darwin-manuscripts-project>. Acesso em: 25 jul. 2020.

DIGITAL HISTORY AND PHILOSOPHY OF SCIENCE. Disponível em: <digitalhps.org>.

HISTORY OF THE MARINE BIOLOGICAL LABORATORY. Disponível em: <history.archives.mbl.edu>. Acesso em: 25 jul. 2020.

HISTORY OF THE MAX PLANCK SOCIETY. Disponível em: <mpiwg-berlin.mpg.de/en/research/projects/DEPT1_458_HistMPS>. Acesso em: 25 jul. 2020.

POINCARÉ CORRESPONDENCE PROJECT. Disponível em: <poincare.univ-nancy2.fr>. Acesso em: 25 jul. 2020.

SOBEK. Disponível em: <sobek.ufrgs.br>. Acesso em: 25 jul. 2020.

REGISTRES DE L'ACADÉMIE. Disponível em: <mpiwg-berlin.mpg.de/en/research/projects/DEPT1_458_HistMPS>. Acesso em: 25 jul. 2020.

RESEARCH COORDINATION NETWORK/NATIONAL SCIENCE FOUNDATION.
Disponível em: <[digitalhps.org/node/184_nsf.gov/awardsearch/
showAward?AWD_ID=1656284&HistoricalAwards=false](https://digitalhps.org/node/184_nsf.gov/awardsearch/showAward?AWD_ID=1656284&HistoricalAwards=false)>. Acesso em:
25 jul. 2020.

SIX DEGREES OF FRANCIS BACON. Disponível em: <[mpiwg-berlin.mpg.
de/en/research/projects/DEPT1_458_HistMPS](https://mpiwg-berlin.mpg.de/en/research/projects/DEPT1_458_HistMPS)>. Acesso em: 25 jul.
2020.

TEXT ENCODING INITIATIVE CONSORTIUM. Disponível em: <tei-c.org>.
Acesso em: 25 jul. 2020.



ARTIGOS

A aliança entre humanidades e tecnologias computacionais e a ressignificação de conhecimento

Isabel Jungk¹

Resumo: As tecnologias informacionais têm transformado as humanidades, especialmente no que tange à pesquisa de dados e produção de conhecimento, que migraram dos suportes analógicos para os digitais, dando surgimento a um campo de estudos interdisciplinares denominado Humanidades Digitais ainda em vias de definição, mas que pode ser compreendido como uma *virada computacional* em todas as áreas das humanidades, pois baseia-se em métodos digitais de processamento de informações. Essa aliança entre humanidades e tecnologias computacionais para produção de conhecimento vêm passando por diferentes etapas de desenvolvimento (primeira, segunda e terceira ondas) e, nessa empreitada, um dos maiores desafios da pesquisa computacionalmente mediada é combinar a busca de informações quantitativas e qualitativas a partir dos dados disponíveis com uma perspectiva crítica, que reflita sobre os valores e premissas fundamentais que orientam tais pesquisas na era da automação computacional. Com base no conceito peirceano de informação, o presente artigo busca mostrar que, a partir dessa materialidade computacional, o conhecimento qualitativo, predominantemente icônico, e o conhecimento quantitativo, predominantemente indicial, se combinam para formar um cabedal simbólico de conhecimento em constante crescimento, sendo influenciados pela ordem simbólica pré-existente, mas sendo capaz também de impactá-la, redimensioná-la e transformá-la, ressignificando o conhecimento em novas sínteses informacionais, aportando novas formas de pensar o próprio ser humano e sua relação com a produção científica e cultural. Parte-se do pressuposto de que essa perspectiva crítica será fundamental para os anos vindouros das Humanidades.

Palavras-chave: Produção de conhecimento. Produção simbólica. Informação qualitativa. Informação quantitativa. Charles S. Peirce.

¹ Isabel Jungk é Doutora em Tecnologias da Inteligência e Design Digital e Mestre em Comunicação e Semiótica pela PUC-SP. cv Lattes: lattes.cnpq.br/2830878145479718. E-mail: isabeljungk@yahoo.com.br.

The alliance between humanities and computational technologies as a process of resignification of knowledge

Abstract: Technologies of information have transformed the humanities, especially with regard to data research and knowledge production, which have migrated from analog to digital media, giving rise to an interdisciplinary field of studies called digital humanities (DH). The DH are still in the process of being defined, but they can be understood as a *computational turn* in all areas of the humanities with its digital methods of information processing. This alliance between the humanities and computer technologies for knowledge production has been going through different stages of development (first, second and third waves). In this endeavor, one of the greatest challenges of computationally mediated research is to combine the search for quantitative and qualitative information from the available data with a critical perspective, reflecting on the fundamental values and premises that guide such research in the era of computational automation. The article adopts C. S. Peirce's concept of information according to which qualitative knowledge is predominantly iconic, whereas quantitative knowledge is predominantly indexical. The two combine to form symbolic knowledge. Symbols grow, and are able to impact, re-dimension and transform knowledge, re-signifying it in new informational syntheses. They provide new ways of thinking about the human mind and its relationship with scientific and cultural production. A critical perspective is fundamental to the future of the humanities.

Keywords: Knowledge production. Symbols. Qualitative information. Quantitative information. Charles S. Peirce.

Humanidades e tecnologias computacionais aliadas na produção de conhecimento

Há 150 anos o conhecimento vem sendo experienciado e praticado, divulgado e apreendido de forma setorizada. As transformações que se cristalizaram nas práticas e disciplinas institucionalizadas que caracterizam as humanidades atualmente são fruto de um processo de setorização e especialização do conhecimento em curso desde a idade média, resultando nas ciências humanas do século XIX, e culminando no edifício epistemológico das humanidades contemporâneas:

Realizado no curso da modernização da universidade medieval, o processo [de subdivisão disciplinar] foi poderosamente moldado pelo surgimento de academias reais nos séculos XVI e XVII, e, no seu rastro, de sociedades eruditas e academias nacionais nos séculos XVIII e XIX. Cada um deles tinha suas próprias licenças sobre o conhecimento, bem como cerimônias, reuniões e publicações profissionais. Na segunda metade do século XIX, com a industrialização em pleno desenvolvimento e a construção de sistemas de ensino público e universidades públicas em curso na Europa e nos Estados Unidos, as humanidades começaram a assumir sua aparência contemporânea. Esta é a era em que a literatura, a filosofia e os estudos clássicos foram separados das ciências naturais e físicas, até mesmo como “história”, e as disciplinas históricas passaram a ser entendidas como expressões de *Wissenschaft*, no duplo sentido de uma “ciência” e uma disciplina dotada de um kit de ferramentas específico para lidar com o registro cultural. (BURDICK *et al.*, 2012, p. 6)

Tal processo de setorização abordava o conhecimento de uma perspectiva positivista, analítica, esmiuçando-o para explorá-lo em profundidade, tendo sido a base do trabalho científico nesse período. A partir da segunda metade do século XX, as desvantagens da excessiva especialização começaram a se fazer sentir, mostrando a necessidade de se estabelecer pesquisas que relacionassem disciplinas e conjugassem conhecimentos de forma múltipla. O movimento acadêmico pela interdisciplinaridade surgiu na França e Itália a partir dos anos 1960, e visava colocar em discussão os critérios fundamentais do ensino nas escolas e universidades. Desde então, muito se argumentou em prol da multi, inter e transdisciplinaridade, mas os esforços no sentido de construir formas integradas de abordar o conhecimento continuaram a esbarrar nas nomenclaturas e estruturas teóricas das tradicionais disciplinas da academia.

O desenvolvimento da cibernética e da linguagem computacional, a partir da década de 1940, tornou possível aos meios computacionais permear as diversas linguagens e aspectos da vida humana. Foi nessa época que as chamadas Humanidades Computacionais surgiram, fruto da aplicação gradual e crescente da computação a bancos de dados textuais. O processo contínuo de expansão tecnológica desembocou na rede mundial de computadores no fim do século passado e no seu crescimento exponencial desde o alvorecer do novo milênio.

Atualmente, as tecnologias digitais embasam todas as mídias e processos comunicacionais, sendo a forma de mediação pervasiva para a produção, pesquisa e divulgação de toda forma de conhecimento na atualidade, o que vêm rompendo e indo além de antigas concepções e estruturas, transformando a universidade de forma inédita. Conforme Schnapp e Presner (2008, §5), “interdisciplinaridade/transdisciplinaridade/multidisciplinaridade são palavras vazias, a menos que impliquem mudanças na linguagem, práticas, métodos e resultados” da academia e, nesse sentido, a atual conjuntura representa um raro momento de transformação cultural, histórica e humana.

As Humanidades Digitais, ou abreviadamente HD, que receberam essa denominação a partir de 2001, representam um verdadeiro marco caracterizado pela oportunidade de crescimento intelectual, semelhante, por um lado, a outros momentos em que a humanidade deu saltos epistemológicos, como a invenção do códex, dos tipos móveis, a descoberta do Novo Mundo e a Revolução Industrial (BURDICK *et al.*, 2012, p. vii), mas que, por outro lado, também traz a possibilidade de ampliação do escopo da investigação científica em cada uma das disciplinas que compõem as humanidades, bem como de sua divulgação e apropriação por parte daqueles que não integram o mundo acadêmico, o que permitirá às humanidades desempenhar um papel criativo a cada dia mais vasto na vida pública.

As HD podem ser consideradas uma abordagem global, transhistórica e transmídia do conhecimento e da construção de significado, uma vez que elas representam uma grande expansão do campo de atuação das humanidades, “precisamente porque elas tratam dos valores, das práticas representacionais e interpretativas, das estratégias de criação de significado, das complexidades e ambiguidades de ser humano em cada campo da experiência e do conhecimento do mundo” (BURDICK *et al.*, *ibid.*).

Humanidades Digitais: campo emergente de estudos

A potencialidade para essa transformação poder ser sentida nos temas enfrentados por aqueles que se dedicam a concretizar, na vida acadêmica, as mudanças engendradas por essa interconexão dinâmica entre poder computacional e humanidades clássicas.

Berry e Fagerjord (2017) apresentam um gráfico estratificado das competências e recursos necessários ao desenvolvimento de projetos em HD, desde os mais fundamentais e invisíveis ao público final, até as publicações e aplicativos que permitem uma perspectiva crítica sobre o conteúdo produzido (Figura 1). Com base na codificação e no pensamento computacional, as instituições de pesquisa montam infraestruturas que incluem métodos, arquivos e metadados digitais, dispostos em estruturas compartilhadas que possibilitam o trabalho de grupos, por vezes numerosos, de pesquisadores e programadores – que trabalham sobre conjuntos de dados de grandes dimensões, cujo tratamento seria inviável a pesquisadores trabalhando individualmente –, e cujo resultado é hospedado em plataformas acessadas via interfaces digitais, que possibilitam sua divulgação, consulta e continuidade de desenvolvimento:

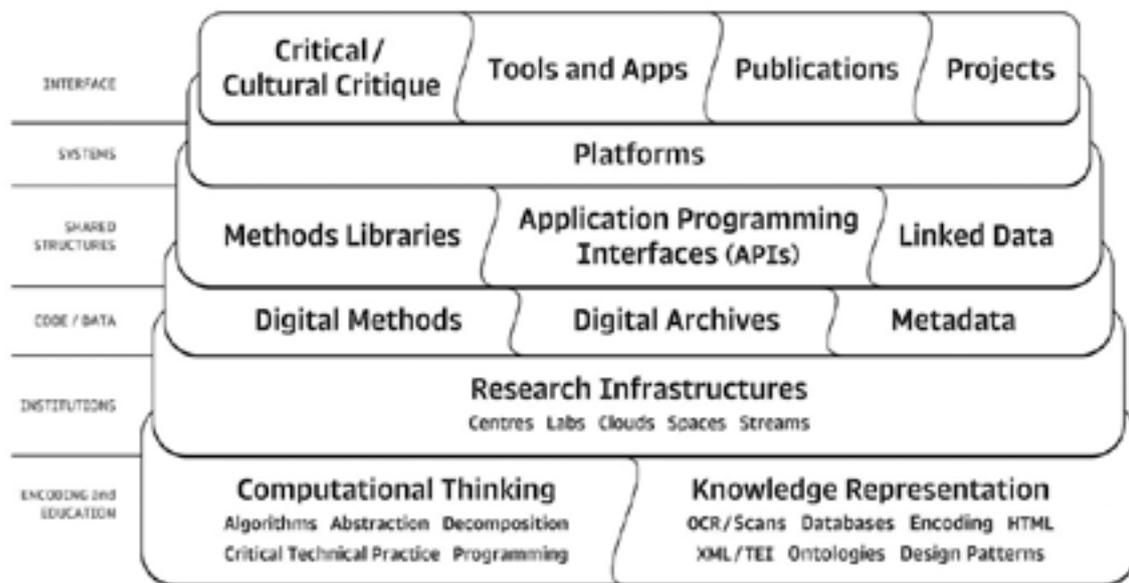


Figura 1. The digital humanities stack (BERRY; FAGERJORD, 2017, p. 19).

Atualmente, tais processos vêm sendo aplicados a todos os campos das artes e ciências sociais, estabelecendo pontes e conexões entre saberes e disciplinas. Os projetos de divulgação de conhecimento e disponibilização de conteúdo são inúmeros – dificilmente uma instituição dedicada ao ensino e/ou cultura não possui um atualmente –, e a quantidade de livros, artigos em revistas especializadas, fóruns de discussões e eventos que têm proliferado nas últimas três décadas sobre as HD é vasto e heterogêneo.

Como observa Gold (2012, p. xii), o campo das HD se move rapidamente, sendo a velocidade e profusão dos debates frequentemente notada com surpresa pelos recém-chegados. As discussões são amplas, mas sempre com vistas à construção de formas integradas de aliar humanidades e sua prática com recursos computacionais: novas metodologias e possibilidades investigativas; práticas colaborativas; requisitos e habilidades dos humanistas digitais; interfaces; ferramentas, linguagens e aplicações digitais; estruturação e compartilhamento de dados; projetos, desenvolvimentos concretos e resultados alcançados; reflexões críticas sobre os valores fundamentais das HD; questionamentos de suas lacunas e vieses; bem como os impactos e as perspectivas futuras das HD na produção e disseminação de conhecimento, dentro e fora da academia. Kirschenbaum traça um retrato da proficuidade de iniciativas no campo das HD:

Os contornos deste aparelho profissional são facilmente detectáveis. Uma organização chamada *Alliance of digital humanities organizations* promove anualmente uma concorrida conferência internacional chamada *Digital humanities*. (Ela surgiu de uma série anual anterior de conferências, organizada em conjunto pela *Association for computers and the humanities* e a *Association for literary and linguistic computing* desde 1989.) Há o livro *Companion to digital humanities*, da Blackwell. Há uma série de livros [...] *Topics in the digital humanities*, da editora da Universidade de Illinois. Existe um periódico acadêmico chamado *Digital Humanities Quarterly*, um dos vários que atendem ao campo, incluindo uma publicação mais recente, *Digital studies/Le champ numérique*, patrocinado pela *Canadian Society for Digital Humanities (Société pour l'Étude des Médias Interactifs)*. A Universidade de Victoria promove anualmente o *Digital Humanities Summer Institute*, para treinar novos acadêmicos. De vital importância, há diversos centros e institutos de Humanidades Digitais (provavelmente pelo menos cem em todo o mundo, alguns deles estabelecidos por uma década ou mais, com equipes que contam com dezenas de integrantes); eles são atendidos por uma organização conhecida como *CenterNet*. Tem havido manifestos de Humanidades Digitais (pelo menos dois) e FAQs [*Frequent Asked Questions*], colóquios e simpósios, workshops e sessões especiais. Sem mencionar [...] as incontáveis vezes que o tema foi visitado em listas de discussão eletrônicas, blogs, perfis do Facebook e feeds do Twitter [...]. (KIRSCHENBAUM, 2012, p. 3-4)

Além dessas, muitas outras publicações, eventos e associações já foram estabelecidas pelo mundo, tais como a *Sociedad Internacional Humanidades Digitales Hispánicas* (HDH), para exemplificar uma iniciativa europeia entre as diversas que existem no continente. Na América Hispânica também há grupos de pesquisa, publicações e eventos dedicados ao tema, como por exemplo no México, Argentina e Colômbia, que já conta com um mestrado específico na área, a *Maestría en Humanidades Digitales* da *Universidad de los Andes*. No Brasil, em outubro de 2013, foi realizado, o primeiro *Seminário Internacional em Humanidades Digitais*, organizado pela Biblioteca Brasileira Guita e José Mindlin da Universidade de São Paulo, por iniciativa do grupo de pesquisa sobre o tema liderado pela professora e pesquisadora Maria Clara Paixão de Sousa, cofundadora da AHDig – *Associação das Humanidades Digitais*. Em 2017, o periódico estadunidense *Digital Humanities Quarterly* publicou uma edição especial com artigos em português, que contou com contribuições de vários pesquisadores brasileiros. Em 2018, foi realizado, no Rio de Janeiro, o *I Congresso Internacional em Humanidades Digitais*, estando a segunda edição do evento programada para 2021.

Para Fitzpatrick (2012, p. 14), “o estado das coisas nas Humanidades Digitais hoje repousa nessa tensão criativa entre disciplinaridade e interdisciplinaridade, entre fazer e interpretar, entre a história do campo e seu futuro”. Todo o desenvolvimento das HD visa a incorporação contínua de novos meios computacionais, fontes e métodos de pesquisa nas humanidades, integrando o analógico e o digital que, em última instância, sempre se refere à interação humana com o mundo, pois, “em um sentido profundo, o digital é sempre fundamentalmente sobre o analógico, ou seja, o humano”, como observa Schnapp (*apud* BIRKLE; DÄWES, 2019, p. 119).

Primeira, segunda e terceira ondas dos estudos em Humanidades Digitais

Segundo Berry e Fagerjord (2017, p. 35), diversas têm sido as tentativas de periodizar as etapas pelas quais vem passando o desenvolvimento das HD (e.g. SCHNAPP; PRESNER, 2009; DAVIDSON, 2012; RAMSEY, 2013, BARNETT *apud* BERRY; FAGERJORD, 2017, *ibid.*). Os autores observam que essas classificações são diferentes modelos para pensar a trajetória dos projetos em HD e compreender seus fundamentos temporais e metodológicos, que continuam sob intenso debate.

Para Berry (2011b, p. 2), as HD podem ser pensadas e apresentadas em três grandes períodos ou ondas. A primeira, desenvolvida entre os anos 1940 e 2001, se convencionou chamar de Computação nas Humanidades (Computing in the Humanities) ou Humanidades Computacionais (Humanities Computing). Nessa fase inicial, predominou a visão das tecnologias em sua função de suporte ao trabalho dos humanistas, bem como de sua aplicação às temáticas das disciplinas humanísticas, ou seja, a tecnologia como mero instrumento de aumento de eficiência, sem nenhum papel crítico (ibid.).

Uma segunda onda, com ênfase nas interfaces e no conhecimento produzido originariamente no formato digital (knowledge born digital) pode ser identificada a partir de 2001, que emergiu com o nome de Humanidades Digitais (*digital humanities*), no esteio do título do livro organizado por Schreibman, Siemens e Unsworth (2001). Conforme Hayles (*apud* BERRY, ibid.), a mudança de denominação sinalizou a elevação do status dos recursos computacionais de mero serviço de suporte a uma empreitada genuinamente intelectual, com práticas profissionais próprias, padrões rigorosos e investigações teóricas autênticas. Os projetos se tornaram maiores e mais complexos, desenvolvendo técnicas computacionais como parte do processo investigativo, que passaram a integrar e ocupar um lugar de destaque no desenvolvimento das humanidades, de modo que “a tecnologia computacional tornou-se fundamentalmente a condição de possibilidade necessária para se pensar sobre muitas das questões levantadas nas ciências humanas hoje” (BERRY, ibid.).

Embora essa classificação em períodos seja discutível, ela pode ser utilizada como ferramenta de análise para refletir sobre as transformações no trabalho das humanidades nas últimas décadas:

A questão de como as Humanidades Digitais realizam suas pesquisas, e se as noções de Humanidades Digitais de primeira e segunda onda capturam o estado atual de diferentes práticas e métodos de trabalho nas Humanidades Digitais permanece contestada. No entanto, esses podem ser conceitos analíticos úteis para pensar sobre as mudanças nas Humanidades Digitais. Podemos, no entanto, observar o seguinte: as Humanidades Digitais de primeira onda envolveram a construção de infraestrutura no estudo de textos de humanidades por meio de repositórios digitais, marcação de texto etc., ao passo que as Humanidades Digitais de segunda onda expandem os limites nocionais do arquivo para incluir obras digitais e, assim, trazer os próprios kits de ferramentas metodológicas das humanidades para examinar materiais originalmente produzidos como digitais [*born digital*], tais como literatura eletrônica (e-lit), ficção interativa (FI), artefatos baseados na web e assim por diante. (BERRY, 2011b, p. 3-4)

Berry (ibid., p. 4) argumenta ainda que, para além da necessidade de explorar a primeira e segunda ondas, também é necessário começar a mapear, mesmo que tentativamente, o que pode ser considerado uma terceira onda das HD, com foco na “*computacionalidade* subjacente das formas contidas em um meio computacional”, e que ele chama de *virada computacional* nas Artes e Humanidades. Isso envolve atentar para o componente *digital* das HD à luz de sua especificidade como meio em si, a fim de pensar de que forma as mudanças midiáticas produzem as mudanças epistêmicas:

Quero sugerir que nem a primeira, nem a segunda onda das Humanidades Digitais realmente problematizaram o que Lakatos (1980) teria chamado de ‘núcleo duro’ das humanidades, [i.e.] as suposições tácitas e as bases ontológicas que sustentam a pesquisa considerada “normal” que os acadêmicos das ciências humanas realizam diuturnamente [...]. De fato, poderíamos dizer que as Humanidades Digitais da terceira onda apontam o caminho pelo qual a tecnologia digital destaca as anomalias geradas em um projeto de pesquisa em humanidades e que leva a um questionamento dos pressupostos implícitos em tal pesquisa, por exemplo, leitura atenta, formação de cânones, periodização, humanismo liberal etc. (BERRY, 2011b, p. 4)

O uso de tecnologias digitais pode assim servir para problematizar onde, como e com quais finalidades as fronteiras disciplinares foram traçadas no passado, especialmente considerando a tendência do digital de dissolver as estruturas institucionais tradicionais (BERRY, 2011a). Berry defende que se passe a ver a importância cultural das possibilidades digitais como a própria ideia unificadora da universidade. Isso significa que, para além do letramento digital e informacional, é necessário pensar sobre o que realmente significa ler e escrever na era computacional, a fim de ter um entendimento crítico sobre o real valor da produção de conteúdo digital e sobre a própria cultura digital em seus aspectos educacionais, caminhando das humanidades do passado, feitas “para alguns humanos”, para as humanidades feitas “para todos os humanos”.

Assim, a visão do desenvolvimento das humanidades em função das possibilidades que os recursos computacionais lhes têm aportado pode ser sintetizada em três etapas: (1ª) Humanidades Computacionais; (2ª) Humanidades Digitais; e (3ª) Humanidades Digitais Críticas. Em termos conceituais, a terceira fase representa o desenvolvimento de um *intelecto digital* em contraposição a uma inteligência digital já disponível, pois a inteligência busca apreender, manipular, reordenar, ajustar, ao passo que o intelecto examina, pondera, questiona, teoriza, critica, contempla, imagina. “A inteligência apreende o significado imediato de uma situação e o avalia. O intelecto avalia as avaliações e procura os significados das situações como um todo” (HOFSTADTER *apud* BERRY, 2011b, p. 8).

Mediação nas Humanidades Digitais: O quantitativo, o qualitativo e o crítico

As três etapas dessa fusão de conhecimento e computacionalidade também podem ser pensadas em termos quantitativos e qualitativos, tanto em relação ao tratamento da informação que empreendem, como à produção do tipo de conhecimento que geram.

Jeffrey Schnapp e Todd Presner, autores de duas versões de um manifesto pelos fundamentos e rumos das HD, *A digital humanities manifesto* (2008) e *Digital humanities manifesto 2.0* (2009), expondo as principais concepções e diretrizes para a geração, transmissão, acessibilidade e preservação do conhecimento das humanidades na era digital, observam que as fases de seu desenvolvimento possuem características específicas. A primeira onda, as Humanidades Computacionais, foi marcada por um caráter predominantemente *quantitativo*, ao digitalizar vastos conjuntos de dados, especialmente textuais, tornando-os processáveis, ou seja, estabelecendo infraestruturas tecnológicas com esse objetivo e usando a capacidade computacional (sistemas de classificação, marcação, codificação de texto e edição acadêmica) para processá-los e analisá-los, inclusive estatisticamente.

A segunda onda tem utilizado as capacidades computacionais para sublinhar o caráter *qualitativo* das informações, provendo novos modos de senti-las, experimentá-las, absorvê-las e interpretá-las, criando os ambientes e o instrumental propício à curadoria do conhecimento, seja ele originariamente digital ou digitalizado, bem como sua divulgação e interação com públicos mais amplos. Segundo Presner (*apud* BERRY 2011b, p. 3), nesta fase, o trabalho das HD “introduz paradigmas disciplinares inteiramente novos, campos convergentes, metodologias híbridas e até mesmo novos modelos de publicação que muitas vezes não são derivados ou limitados à cultura impressa”. Essas etapas não se esgotam nos períodos em que se mostraram predominantes, e os aspectos quantitativos e qualitativos da pesquisa humanística podem ser observados em constante permuta e desenvolvimento:

A primeira onda de trabalho das Humanidades Digitais foi quantitativa, mobilizando os poderes de pesquisa e recuperação do banco de dados, automatizando a linguística de corpus, arranjando *hypercards* em feixes críticos. A segunda onda é de caráter qualitativo, interpretativo, experiencial, emotivo e generativo. Ele aproveita os kits de ferramentas digitais a serviço dos principais pontos fortes metodológicos das Humanidades: atenção à complexidade, especificidade do meio, contexto histórico, profundidade analítica, crítica e interpretação.

Essa dicotomia cruamente esboçada não exclui a potencialidade emocional, mesmo sublime, do quantitativo, nem exclui os encaixes da análise quantitativa em estruturas qualitativas. Pelo contrário, ela imagina novos acoplamentos e escalas que são facilitados tanto por novos modelos de prática de pesquisa quanto pela disponibilidade de novas ferramentas e tecnologias. (SCHNAPP; PRESNER, 2009, §II)

De modo geral, pode-se observar que a primeira fase foi marcada pela exploração da computação para o processamento da informação, ao passo que a segunda vem explorando a experiência que a tecnologia pode prover da informação, para transformá-la em conhecimento. Esses dois vetores vêm convergindo para o que pode ser considerada uma terceira etapa, a das Humanidades Digitais Críticas que, em um movimento autorreflexivo, se voltam para o exame minucioso de sua materialidade, isto é, das formas de *mediação computacional* que possibilitam que o conhecimento seja produzido por meios digitais.

O trabalho em HD se baseia na plasticidade das formas digitais, isto é, em sua maleabilidade em função de seu substrato comum, a linguagem computacional que as permeia, bem como em suas premissas e estruturas codificadas que constituem a forma como elas apontam para uma nova maneira de trabalhar com a representação e a mediação:

Mediar um objeto, um dispositivo digital ou *computacional* requer que este objeto seja traduzido no código digital que ele pode entender. Esta transformação mínima é efetuada através do mecanismo de entrada de um dispositivo sociotécnico dentro do qual um modelo ou imagem é estabilizado e processado. Ele é então transformado internamente, dependendo de uma série de intervenções, processos ou filtros e, eventualmente, é exibido ao final como resultado de um cálculo, geralmente em um formato visual. [...] O ponto principal é que sem a possibilidade da codificação *discreta*, não há objeto para o dispositivo computacional processar. (BERRY, 2011b, p. 1-2)

Toda forma de codificação pode ser compreendida como um processo semiótico de mediação. Para Peirce (EP 2, p. 544, 1906), um signo é uma espécie de meio de comunicação e, “como meio, o Signo está essencialmente em uma relação triádica, com seu Objeto, que o determina, e com seu Interpretante, que ele determina”. Isso equivale a dizer que um signo é algo que faz a mediação entre o objeto ou objetos ao qual se refere e um interpretante ou interpretantes que resultam dessa mediação, de modo que ambos, objeto e interpretante se entrelaçam e especificam mutuamente, por meio da mediação desse signo (PEIRCE, EP 2, p. 410, 1907). Em termos computacionais, essa relação triádica une dinamicamente:

- (i) uma representação ou signo em si mesmo: a informação codificada em determinada linguagem, capaz de ser processada em determinado(s) software(s)/aplicativos(s);
- (ii) um objeto ou referente: o objeto mediado, que tem parte de seus atributos reais representados pela informação codificada; e
- (iii) um resultado mediatório ou interpretante: entendido como o produto da capacidade constitutiva do signo de se fazer representar em novas formas de significação em referência ao objeto representado, ou seja, é o *output* ou resultado buscado com o processamento da informação codificada via determinado software ou aplicativo.

Dessa forma, os signos do universo computacional são produzidos por meio das linguagens que codificam as informações sobre entidades, objetos e fenômenos reais do mundo, tornando possível inseri-los no fluxo eletrônico da máquina, representando-os, ainda que sempre sob determinados aspectos e nunca de forma completa, uma vez que somente a entidade representada pode conter nela mesma todos os atributos e, portanto, todas as informações que a constituem.

Ao fragmentar o mundo em porções discretas de informação, uma parte dessa informação é necessariamente perdida, descartada, para possibilitar ao sistema computacional armazená-la como representação codificada e subsequentemente processá-la via procedimentos algorítmicos, a fim de produzir novos conhecimentos – novos interpretantes – sobre determinada entidade, objeto ou fenômeno. Esses interpretantes são novos signos, que continuam o processo de semiose ou significação dos objetos sob investigação ou visualização.

Por sua vez, a codificação das informações disponíveis a respeito do mundo resulta igualmente, conforme Berry (2011b, p. 2), em outras “situações do mundo real onde a computação é orientada por eventos² e dividida em processos discretos para realizar uma tarefa específica buscada pelo usuário”, possibilitando a manipulação técnica da realidade e impactando-a de acordo com seus métodos, estruturas e programação, pontos cruciais das HD que é necessário manter constantemente sob escrutínio crítico. Para o autor,

² A programação orientada a eventos (*event-driven programming*) é um paradigma de programação pelo qual o fluxo de execução do programa é determinado por certas ações do usuário ou de outros programas, chamadas eventos, tais como um clique do mouse, o acionamento de funções através de teclas ou via sistema operacional, entre outras.

um computador exige que tudo seja transformado do fluxo contínuo de nossa realidade cotidiana em uma grade de números que podem ser armazenados como uma representação da realidade que pode, em seguida, ser manipulada usando algoritmos. Esses métodos subtrativos de compreensão da realidade (*episteme*) produzem novos conhecimentos e métodos para o controle da realidade (*techne*). Eles fazem isso por meio de uma mediação digital, que as Humanidades Digitais estão começando a tomar seriamente como sua problemática. (BERRY, *ibid.*)

A esse processo de mediação, Berry (*ibid.*) atribui a característica de *estratificação digital* da realidade (*digital folding of reality*), por meio da qual ele defende que é possível abordar a cultura e seus meios de produção de forma radicalmente nova. Dito de outra maneira, o problema das HD nesta terceira fase crítica e reflexiva é uma *questão mediatória*, uma interrogação sobre como as estruturas e ontologias³ computacionais representam certas porções de conhecimento acerca da realidade e quais são as diretrizes, objetivos e vieses que as embasaram, tenha sido de forma tácita ou explícita e propositada. Essa perspectiva crítica possibilitará a reflexão sobre os resultados e consequências em diversos níveis desses processos tecnológicos de produção epistemológica e mediação do conhecimento.

Materialidade computacional e informação

Na transformação que as tecnologias informacionais vêm impondo à pesquisa e produção de conhecimento nas humanidades, cada forma concreta de corporificar a informação carrega as potencialidades próprias à materialidade de que está constituída. Para Schnapp (*apud* BIRKLE; DÄWES, 2019, p. 119), é um engano traçar delineamentos excessivamente nítidos e mesmo qualquer oposição entre o digital e o analógico, atribuindo materialidade ao analógico e imaterialidade ao digital.

Conforme Munster (*apud* RYAN *et al.*, 2014, p. 328), “a materialidade não pode ser erroneamente tomada como uma simples qualidade ou propriedade das tecnologias digitais. Ao invés disso, ela pode ser compreendida como uma abordagem expandida para analisar questões e áreas na computação e na cultura”. A materialidade computacional, para além das unidades binárias de 0 e 1 que constituem sua propriedade eletrônica (física) mais elementar, se traduz pela condensação das forças e relações que constituem tanto hardware como software, e que são mobilizadas na direção de certos objetivos pelo pesquisador.

³ De modo geral, em computação, uma ontologia compreende a classificação das informações através de uma representação, seja ela uma nomeação formal ou descrição/definição de acordo com um sistema de categorias que classifica suas propriedades e relações interdependentes para fins de processamento.

Isso permite observar que as forças epistemológicas condensadas na materialidade computacional utilizadas pelas HD privilegiaram, de acordo com suas possibilidades em cada etapa de seu desenvolvimento, aspectos diferentes da informação. Primeiramente, a abordagem numérica prevaleceu, isto é, o cálculo com base em métodos quantitativos foi utilizado para processar as informações extraídas da materialidade física (analogicamente armazenada) que originalmente servia de suporte ao conhecimento. A partir do momento em que o cabedal de conhecimento humanístico se tornou legível e disponível para o processamento algorítmico, os métodos das ciências humanas se aproximaram dos métodos quantitativos das ciências naturais.

Com a evolução dos meios digitais, que passaram a disponibilizar novas interfaces gráficas, tornou-se possível a visualização didática do conhecimento para fins pedagógicos e de divulgação, bem como sua abordagem experiencial para fins de vivência do conhecimento, favorecendo sua exploração sensorial e qualitativa por meio de imagens, gráficos, vídeos, simulações, ambientes virtuais, entre outros. Apesar da complementaridade e convergência dessas abordagens, ainda

permanece um coro de vozes contemporâneas lamentando outra crise “definitiva” na cultura humanística, mais um sacrifício da qualidade no altar da “mera” quantidade. [...] Não acreditamos que as humanidades estejam em perpétua crise ou ameaçadas por outra batalha por legitimidade com as ciências. Em vez disso, vemos este momento como marcando uma mudança fundamental na percepção das atividades criativas fundamentais do ser humano, em que os valores e conhecimentos das humanidades são vistos como cruciais para moldar todos os domínios da cultura e da sociedade. (BURDICK *et al.*, 2012, p. vii)

Essa combinação e convergência entre quantitativo e qualitativo e sua importância para as atividades criativas e investigativas e os valores que as conduzem encontra ressonância no conceito peirceano de informação.

O crescimento da informação: o icônico, o indexical e o simbólico

Para Peirce, um símbolo condensa informação denotativa e conotativa, isto é, quantitativa e qualitativa, que se combinam no conhecimento que o símbolo é capaz de veicular. Tal informação pode ser logicamente definida como a conjugação da amplitude de objetos denotados pelo símbolo e da profundidade de atributos significados. Conforme Nöth e Amaral,

Denotação, referência, extensão ou largura são conceitos que se referem aos objetos do conhecimento empírico aos quais os símbolos se aplicam. Significação, sentido, intenção, significado, compreensão ou profundidade são termos que se referem ao conhecimento verbal associado aos símbolos [...]. Nos primeiros escritos sobre a teoria da informação, essas duas dimensões lógico-semânticas são apresentadas com os nomes de largura e profundidade e, nos seus escritos mais tardios, Peirce utiliza os termos denotação e significação [...]. Numa definição lógica mais precisa, a denotação de um símbolo é o conjunto de todas as “coisas reais” das quais ele pode ser predicado [...] e a significação de um símbolo, por sua vez, é o conjunto dos predicados atribuíveis ao sujeito numa proposição verdadeira. (NÖTH; AMARAL, 2011, p. 8)

A informação contida nos símbolos, isto é, nos conceitos, proposições e teorias a respeito do mundo, é uma combinação de diferentes tipos de signos atuantes no processo de semiose ou significação da realidade, sempre complexa. Signos podem apresentar seus objetos de formas icônicas, por meio de semelhanças qualitativas entre suas aparências, estruturas e significados (imagens, diagramas, analogias, metáforas etc.), ou podem se referir a eles de formas indexicais (rastros, sinais, índices, indexadores, nomes etc.), por meio de relações dinâmicas de contiguidade no tempo e espaço, ou mediá-lo de formas simbólicas, por meio de representações convencionais que podem expressar hábitos culturais ou naturais (regras, normas, leis, relações de causalidade etc.), incluindo componentes icônicos e indexicais em sua significação. Em termos peirceanos, a relação entre esses aspectos no símbolo é representada pela fórmula: *Informação = Significação X Denotação*, de maneira que

Tanto a denotação como a significação podem ser determinados por suas quantidades relativas. A quantidade de significação contida num termo cresce com o número de predicados atribuídos a este termo enquanto que a quantidade de denotação de um termo (sua extensão) cresce com o número de coisas reais que o termo representa. Dentre essas duas quantidades, há uma relação de proporcionalidade inversa [...]: qualquer adição à significação de um termo diminui o escopo de sua denotação [...]. Assim, quanto mais significação um símbolo contiver, menor será o número de coisas reais por ele denotadas e vice-versa. (NÖTH; AMARAL, 2011, p. 8)

Essa proporcionalidade é aplicável a termos, isto é, palavras e conceitos que, quanto mais gerais e amplos, abrangem uma maior quantidade de objetos por possuir uma alta denotação e uma baixa conotação. Porém, quanto mais específicos, maior sua conotação, pois expressam um maior número de atributos de seu objeto, embora, por isso mesmo, denotam uma quantidade menor de objetos. Isso pode ser demonstrado de maneira simples, tomando como exemplo a palavra “árvore”, símbolo

que possui ampla denotação, pois se refere a todos os tipos de árvores, e baixa conotação em função de sua alta generalidade, pois contém poucos elementos característicos atribuíveis a seu objeto. À medida que se especifica o tipo de árvore, os elementos conotativos aumentarão, diminuindo a quantidade de objetos denotados (cf. *ibid.*, p. 8-9).

Conforme Brock (1997, p. 560), em Peirce, ícone, índice, símbolo são definidos em termos de denotação ou extensão e conotação ou compreensão, de forma que ícones conotam sem denotar, índices denotam sem conotar, e símbolos ou termos gerais denotam em virtude de conotar (*ibid.*). Isso mostra porque o símbolo funciona como uma síntese dessas formas de significação, pois a informação quantitativa, predominantemente indicial, contribui para a denotação mais precisa do objeto representado, ao passo que a informação qualitativa favorece sua conotação. A informação veiculada resulta assim da combinação desses aspectos, pois o símbolo contém elementos de iconicidade e indexicalidade ou indicialidade, que Peirce considerou serem as propriedades semióticas do símbolo, “pois a denotação, extensão ou aplicação, isto é, o poder aplicativo, referencial do símbolo corresponde ao seu ingrediente indicial, a significação, conotação ou profundidade corresponde ao seu ingrediente icônico”, como observa Santaella (2001, p. 268).

Embora essa proporcionalidade inversa seja aplicável a símbolos isoladamente considerados, o patrimônio simbólico humano tende sempre a crescer. Nesse sentido, pode-se observar que em todas as fases das HD, como campo interdisciplinar que é, foram sendo produzidos novos conjuntos de informação, cuja significação, ora como maior conteúdo indicial, ora com maior conteúdo icônico, sempre sintetizou novos conhecimentos em formas simbólicas de pensamento. Tais aspectos semióticos são observados igualmente por outros autores. Sobre o aspecto indicial, Berry observa que compreender os processos de codificação e as linguagens computacionais que permitem o desenvolvimento das investigações e projetos em HD pode ser uma rica maneira de compreender a produção cultural de forma mais geral:

Há uma dimensão cultural inegável na computação e nos recursos mediais do software. Essa conexão aponta para a importância de se envolver com e compreender o código de computador; de fato, o código de computador pode servir como um índice de cultura de forma mais geral (imaginem as Humanidades Digitais mapeando diferentes linguagens de programação a respeito das possibilidades culturais e práticas que elas oferecem, por exemplo, HTML para cibercultura, AJAX para mídias sociais etc.), sem mencionar o mapeamento da ‘edição’ de softwares para novas formas de narrativa cinematográfica, música e arte em geral, ou crítica cultural por meio das Humanidades Digitais. (BERRY, 2011b, p. 5)

Por outro lado, como observam Schnapp e Presner (2008, §§10-21), a escala da complexidade do real ultrapassa as capacidades de investigação pessoais, necessitando de times de pesquisadores para ser processada e mediada por novos modelos de produção e reprodução do conhecimento humanístico, sendo esse o desafio contemporâneo das humanidades: construir modelos de criação e compartilhamento do conhecimento que façam frente a essa realidade cada vez mais ampla e distribuída.

Em termos simbólicos, é possível observar que a revolução que as HD promovem se baseia em modos mais abrangentes de *integração* do conhecimento, isto é, na construção de representações mais amplas a partir dos mosaicos de conhecimento humanístico especializado, inclusive ampliando seu alcance a públicos transuniversitários. As HD representam, assim, a realização entretecida da pesquisa no século XXI: uma triangulação do fazer artístico, do exercício do comentário-crítica e do alcance da divulgação, fundindo investigação acadêmica, pedagogia, publicação e prática (ibid.).

Nesse cenário promissor, o conhecimento qualitativo, predominantemente icônico, e o conhecimento quantitativo, predominantemente inercial, se combinam para formar um cabedal simbólico de conhecimento, sempre em crescimento. Como afirma Peirce (EP 2, p. 10, 1894; EP 2, p. 263, 1903), o símbolo é algo vivo: *Omne symbolum de symbolo*. Símbolos surgem por desenvolvimento a partir de outros signos e seu significado se expande invariavelmente, incorporando novos elementos e deitando fora elementos antigos, disseminando-se, ampliando seu significado por meio do uso, da experiência e de outros símbolos. O surgimento de novos símbolos é influenciado pela ordem simbólica pré-existente, mas estes são igualmente capazes de impactá-la, redimensioná-la e transformá-la, ressignificando o conhecimento, aportando novas formas de pensar o próprio ser humano e sua relação com a produção científica e cultural.

Ressignificação de conhecimento em novas sínteses informacionais

Tradicionalmente, as humanidades se dedicam às palavras, aos objetos e às imagens de eras passadas e do presente e, como observam Berry e Fagerjord (2017, p. 114), “em uma era pós-digital, eles estão em formatos digitais, não somente os textos fundamentais, imagens e achados arqueológicos que são representados nos arquivos digitais, mas também os textos acadêmicos que os descrevem e discutem”. Essa mediação codificada de os todos suportes do conhecimento passado e da produção atual de conhecimento representa uma *virada epistemológica*, com transformações individuais, institucionais e conceituais profundas.

No nível individual, a tecnologia estende a capacidade mental humana, conferindo-lhe ampliadas possibilidades de manipular o conhecimento adquirido, e de utilizar e expandir sua imaginação e raciocínio, com vistas à produção científica e cultural de novos conteúdos, sem dúvida, mas igualmente visando a ampliação do alcance de valores humanísticos de igualdade e sabedoria, processo evolutivo ainda em vias de concretização:

Contar com a tecnologia de uma forma mais radicalmente descentrada, depender de dispositivos técnicos para preencher as lacunas de nossas mentes e conectar o conhecimento de novas maneiras, mudará nosso entendimento do conhecimento, da sabedoria e da inteligência em si mesmas. [...] As técnicas computacionais podem nos dar maiores poderes de pensamento, maior alcance para nossa imaginação e, possivelmente, permitir nos reconectarmos às noções políticas de igualdade e redistribuição com base no potencial da computação para dar a cada um de acordo com suas necessidades e habilidades. (BERRY, 2011b, p. 10)

A era da produção de conhecimento digitalizado é *colaborativa*, pois necessita do trabalho realizado coletivamente, é *cumulativa*, pois a vertiginosa capacidade de memória computacional permite o armazenamento e processamento de praticamente toda a informação gerada, e, ainda, é *iterativa*, uma vez que a própria produção potencializa suas realizações, abrangendo o digital e o físico de forma inseparável, sendo capaz de alcançar públicos cada vez mais heterogêneos em função de sua plasticidade e capacidade multifacetada de manipulação e apresentação.

Contudo, conforme Berry (2011b, p. 8), se por um lado, os conjuntos digitais de informação que estão sendo construídos prometem grandes mudanças no nível do ator humano individual – inclusive impactando sua autonomia em relação à absorção e produção de conhecimento –, as vastas quantidades de dados e sua ampla disponibilidade também podem ser desestabilizadoras, causando certa desorientação social, que precisará ser mediada através de racionalidades computacionais que passem pelo crivo crítico da razão, e que podem ser entendidas como *modos supercríticos de pensamento* possibilitados pelo acoplamento da inteligência humana à inteligência artificial na era computacional:

A tecnologia possibilita o acesso aos bancos de dados do conhecimento humano de qualquer lugar, desconsiderando e contornando os tradicionais guardiões do conhecimento no estado, nas universidades e no mercado. Não parece mais ser o professor que lhe diz o que você deve procurar e os ‘três argumentos a favor’ e os ‘três argumentos contra’ algo. Isso introduz não apenas um momento de desorientação social, com indivíduos e instituições inundadas de informações, mas também oferece uma solução computacional para esse estado

de eventos na forma de racionalidades computacionais – algo que Turing (1950) descreveu como modos supercríticos de pensamento. Ambas as forças são sustentadas em um nível estrutural profundo pelas condições de possibilidade sugeridas pelo código de computador. (BERRY, 2011b, p. 8)

As facetas quantitativas, qualitativas e reflexivas do conhecimento podem ser exploradas computacionalmente e recombinadas de maneira crítica em uma nova fase das HD. Tais possibilidades investigativas encontram seu fundamento no potencial constitutivo do signo de produzir sempre novas significações de maneira autônoma. Esse potencial nunca se esgota, uma vez que a maneira como o símbolo representa seu objeto, isto é, a parcela de informação que ele é capaz de veicular, faz a mediação entre entidades dotadas igualmente de realidade própria, que se constituem em seus objetos, e que contém atributos parcialmente explorados ou ainda inexplorados, que sempre podem ser explorados a partir de outras perspectivas e pontos de vista, ou com outros objetivos, gerando novos resultados interpretativos a cada iteração, que podem ser obtidos de diversas formas e em variados níveis.

Como ressalta Berry (2019, p. 64), o problema mais específico das HD é ver sua contribuição à cultura apenas como processamento de sinais digitais, ao invés de vê-lo como processamento de símbolos. Essa diferenciação é crucial, pois, entendidas como um conjunto simbólico, as tecnologias utilizadas pelas humanidades têm grandes implicações para os modelos humanísticos de pesquisa, perspectiva evidenciada pela terceira etapa crítica das HD, que reflete sobre a produção, disseminação e consumo de conhecimento na era digital.

A academia, as instituições que detêm os meios de armazenamento e divulgação de conhecimento, e a sociedade como um todo precisam estar atentas a que, para conhecer em um mundo permeado por tecnologias, é preciso ir além de plataformas, protocolos, ferramentas e linguagens computacionais e se engajar na corporificação de princípios e valores teóricos antigos e novos advindos das humanidades.

Assim, a potência da materialidade computacional e as racionalidades que a constituem favorecem representar a complexidade do real a partir de modelos igualmente mais complexos, capazes de refletir a realidade com mais precisão e acuidade no presente, ressignificando e ampliando o conhecimento passado, para engendrar novas sínteses epistemológicas que possibilitem manipular a concretude das realizações humanas em prol da própria humanidade, o que constitui novas possibilidades de um novo humanismo praticado com base na revolução digital.

Considerações finais: automação computacional e humanidades

Ao serem digitalizados, diferentes conteúdos se transformam em dados até certo ponto homogêneos, que se tornam igualmente comparáveis e processáveis computacionalmente. Sempre houve uma camada física sob a cultura simbolicamente expressa, contudo a materialidade digital torna intercambiável a informação contida nos diversos sistemas digitais, arquivos e interfaces, característica que responderá pelos próximos saltos epistemológicos a serem dados.

Para Berry (2011b, p. 12-13), essas vastas quantidades de informação estruturada abrem novas formas de estudar e compreender o fenômeno humano, engendrando um sistema coletivo de informação, que pode ser manipulado para gerar os *outputs* desejados. Contudo, os modos de codificação precisam ser compreendidos para serem *decodificados em termos críticos* pelo pesquisador, para que os métodos criados e utilizados sejam capazes de impactar a ordem simbólica, indo além da mera utilização instrumental das informações. Daí a importância de uma visão reflexiva sobre as técnicas computacionais e sua lógica, os objetivos que norteiam sua criação e seus potenciais efeitos:

Porém, menos entendida é a maneira pela qual os arquivos digitais que estão sendo criados são profundamente computacionais em sua estrutura e conteúdo, porque a lógica computacional está enredada com as representações digitais de objetos físicos, textos e artefatos “originariamente digitais” [*born digital*]. Técnicas computacionais não são meramente um instrumento manejado por métodos tradicionais; em vez disso, eles têm efeitos profundos em todos os aspectos das disciplinas. Não só introduzem novos métodos, que tendem a focar a identificação de novos padrões nos dados em contraposição com o princípio da narrativa e do entendimento [...]. (BERRY, 2011b, p. 13)

A automação computacional revoluciona e alavanca o poder de processar informações com amplitude, profundidade e escala sem precedentes, o que tem consequências para todos os aspectos da vida em sociedade:

Usamos o termo “automação computacional” para nos referir à extensão atual da automação: a incorporação do aprendizado de máquina e da análise de dados em grande escala na produção econômica, na academia e na vida cotidiana. A automação computacional está revolucionando os empreendimentos humanos, científicos, culturais e artísticos, que estão constantemente sendo reorganizados em novas estruturas temporais, espaciais e conceituais com diferentes graus de agência e reflexividade. (BERRY *et al.*, 2019, p. 61)

Há muito tempo existem grandes quantidades de dados disponíveis, o que não havia eram formas eficientes e rápidas de extrair informação útil a partir deles (o atual *big data*), o que o pesquisador tradicional fazia de maneira manual, analógica, sempre lenta e limitada pela escala das possibilidades humanas. Igualmente, sempre houve metadados envolvidos nas atividades de pesquisa, no entanto sua visualização e processamento não era tão estruturada como acontece nas atuais ontologias computacionais. O crescimento do alcance da automação computacional será contínuo e, com a incorporação da aprendizagem de máquina (*machine learning*), a cada passo de seu desenvolvimento, será capaz de revelar novos padrões, regularidades e características do estar do homem no mundo.

Contudo, o valor cultural da automação computacional nas humanidades precisará ser ponderado para além da alta performance quantitativa, pois novos métodos criados a partir das tecnologias digitais serão igualmente novas formas de lidar com a realidade, modificando o modo como a entendemos, vivenciamos e transformamos. Os métodos investigativos também serão cada vez mais permeados pela mediação computacional, cujas ferramentas nem sempre são desenhadas especificamente para o trabalho em humanidades, o que precisa ser levado em consideração pelo pesquisador em seus processos de investigação, sejam eles indutivos, dedutivos ou abduativos. Consequentemente, analisar os valores que permeiam pesquisas e projetos e questionar as premissas fundamentais que os orientam, identificando possíveis vieses em suas buscas e resultados será cada vez mais a tarefa primordial do pesquisador.

É urgente compreender os limites das técnicas computacionais disponíveis para pensar sua aplicação e ir além delas, tomando decisões interpretativas e sugerindo modificações e mudanças de rota quando necessário, sem permitir que dados, modelagem e visualizações eclipses a teoria e a crítica das quais podem e devem surgir a quebra dos próprios paradigmas que a computação no seu estágio atual também impõe.

As aptidões computacionais possibilitam o hibridismo científico, expandindo fronteiras interdisciplinarmente, unificando campos e áreas do saber, em um impulso permanente rumo a uma revolução em todas as formas de conhecimento, oferecendo novas modalidades para sua representação. Como observou Reategui (2019), com o tempo, as humanidades provavelmente não precisarão mais de uma denominação específica, pois toda pesquisa incorporará a dimensão computacional, que será um de seus pilares. Enfrentar esses desafios investigativos e críticos certamente desvelará novos caminhos epistemológicos a trilhar, e tal perspectiva será fundamental para os anos vindouros das humanidades.

Referências

BIRKLE, Carmen; DÄWES, Birgit. “Old media don’t go away, they mutate”: An interview with Jeffrey Schnapp. *Amerikastudien/American Studies*, n. 64.1, p. 111-125, January, 2019. Disponível em: <amst.winter-verlag.de/article/amst/2019/1/10>. Acesso em: 14 maio 2020.

BERRY, David M. Digital humanities: First, second and third wave. *Stunlaw: Philosophy and Critique for a Digital Age*, Blog. Postado em 14/01/2011a. Disponível em: <stunlaw.blogspot.com/2011/01/digital-humanities-first-second-and.html>. Acesso em: 14 maio 2020.

BERRY, David M. The computational turn: Thinking about the digital humanities. *Culture Machine*, 12. Open Humanities Press, 2011b. Disponível em: <n9.cl/pdmt>. Acesso em: 14 maio 2020.

BERRY, David M.; FAZI, M. Beatrice; ROBERTS, Ben; WEBB, Alban. No signal without symbol: Decoding the digital humanities. *Debates in the digital humanities* 2019. Matthew K. Gold & Lauren F. Klein (eds.) Minneapolis: University of Minnesota Press, 2019. Disponível em: <jstor.org/stable/10.5749/j.ctvg251hk.8> Acesso em: 14 maio 2020.

BERRY, David M.; FAGERJORD, Anders. *Digital humanities: Knowledge and critique in a digital age*. Cambridge: Polity Press, 2017.

BROCK, Jarret. The development of Peirce’s theories of proper names. *Studies in the Logic of Charles Sanders Peirce*. Nathan Houser (ed.). Bloomington: Indiana University Press, 1997.

BURDICK, Anne; DRUCKER, Johanna; LUNENFELD, Peter; PRESNER, Todd; SCHNAPP, Jeffrey. *Digital humanities*. Cambridge, MA: MIT Press, 2012.

FITZPATRICK, Kathleen. The Humanities, done digitally. *Debates in the digital humanities*. Matthew K. Gold (ed.). Minneapolis, MN: University of Minnesota Press, 2012.

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

GOLD, Matthew K.; KLEIN, Lauren F. (eds.). *Debates in the digital humanities* 2019. Minneapolis, MN: University of Minnesota Press, 2019.

JUNGK, Isabel. Métodos para a investigação do real. *Teccogs: Revista Digital de Tecnologias Cognitivas*, TIDD-PUC-SP, São Paulo, n. 12, p. 37-65, jul-dez. 2015. <pucsp.br/pos/tidd/teccogs/artigos/2015/edicao_12/teccogs12_artigo01.pdf>. Acesso em: 14 maio 2020.

KIRSCHENBAUM, Matthew. What is digital humanities and what's it doing in English Departments? *Debates in the digital humanities*. Matthew K. Gold (ed.). Minneapolis, MN: University of Minnesota Press, 2012.

MARQUES, Fabrício. A realidade que emerge da avalanche de dados. *Revista Pesquisa Fapesp*, n. 255, maio, 2017. Disponível em: <revistapesquisa.fapesp.br/2017/05/23/a-realidade-que-emerge-da-avalanche-de-dados/>. Acesso em: 14 maio 2020.

NÖTH, Winfried; AMARAL, Gustavo Rick. A teoria da informação de Charles S. Peirce. *Teccogs: Revista Digital de Tecnologias Cognitivas*, TIDD-PUC-SP, São Paulo, n. 5, p. 4-29, 2011. Disponível em: <pucsp.br/pos/tidd/teccogs/artigos/2011/edicao_5/1-teoria_da_informacao_de_charles_sanders_peirce-winfried_noth-amaral_gurick.pdf>. Acesso em: 14 maio 2020.

PEIRCE, Charles S. *The Essential Peirce: Selected Philosophical Writings*. Vol. 1 (1863-1893). Houser, N.; Kloesel, C. (eds.), 1992. Vol. 2 (1893-1913). The Peirce Edition Project (ed.), 1998. Bloomington, IN: Indiana University Press, 1992-1998.

REATEGUI, Eliseo. Contribuições da Inteligência Artificial para as Humanidades Digitais. Palestra proferida em 14 ago. 2019 no Programa de Estudos Pós-graduados em Tecnologias da Inteligência e Design Digital – PUC-SP, 2019.

RODRÍGUEZ-ORTEGA, Nuria. Cinco ejes para pensar las Humanidades Digitales como proyecto de un nuevo humanismo digital. *Artnodes*, n. 22, p. 1-6, UOC, 2018. Disponível em: <[dx.doi.org/10.7238/a.voi22.3263](https://doi.org/10.7238/a.voi22.3263)>. Acesso em: 14 maio 2020.

RYAN, Marie-Laure; EMERSON, Lori; ROBERTSON, Benjamin J. *The Johns Hopkins Guide to Digital Media*. Baltimore: Johns Hopkins University Press, 2014.

SANTAELLA, Lucia. *Matrizes da linguagem e pensamento: sonora, visual, verbal: aplicações na hipermídia*. 3ª ed. SP: Iluminuras: FAPESP, 2001.

SCHNAPP, Jeffrey; PRESNER, Todd. A digital humanities Manifesto. University of California, Los Angeles, *UCLA Website*, postado em 15 dez. 2008. Disponível em: <manifesto.humanities.ucla.edu/2008/12/15/digital-humanities-manifesto/>. Acesso em: 14 maio 2020.

SCHNAPP, Jeffrey; PRESNER, Todd. Digital humanities manifesto 2.0. University of California, Los Angeles, *UCLA Website*, postado em 29/05/2009. Disponível em: <manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20/>. Acesso em: 14 maio 2020.

SCHREIBMAN, Susan; SIEMENS, Ray; UNSWORTH, John (eds.). *A new companion to digital humanities*. Malden: Wiley Blackwell, 2016.

SCHREIBMAN, Susan; SIEMENS, Ray; UNSWORTH, John (eds.). *Companion to digital humanities*. Malden: Wiley Blackwell, 2001.

Um breve guia para as Humanidades Digitais¹

Anne Burdick²

Johanna Drucker³

Peter Lunenfeld⁴

Todd Presner⁵

Jeffrey Schnapp⁶

Tradução do inglês: Isabel Jungk⁷

Resumo: À medida que metodologias, ferramentas e habilidades digitais tornam-se crescentemente centrais para trabalhar nas humanidades, questões concernentes aos fundamentos, resultados de projetos, avaliações e design se tornaram urgentes. As especificações proporcionam um conjunto de critérios para guiar aqueles que realmente trabalham nas Humanidades Digitais, assim como aqueles a quem se pede para avaliar e financiar pesquisadores, projetos e iniciativas das Humanidades Digitais.

Palavras-chave: Humanidades Digitais. Metodologias. Projetos. Avaliação.

¹ Agradecimentos às autoras e aos autores de *Digital humanities* (MIT Press, 2012) pela autorização desta tradução do capítulo *Short guide to the digital humanities* (licença CC).

² E-mail: burdick@artcenter.edu.

³ E-mail: drucker@gseis.ucla.edu.

⁴ E-mail: lunenfeld@arts.ucla.edu.

⁵ E-mail: presner@humnet.ucla.edu.

⁶ E-mail: jeffrey@metalab.harvard.edu.

⁷ Isabel Jungk é Doutora em Tecnologias da Inteligência e Design Digital e Mestre em Comunicação e Semiótica pela PUC-SP. CV Lattes: lattes.cnpq.br/2830878145479718. E-mail: isabeljungk@yahoo.com.br.

Short guide to the digital humanities

Abstract: As digital methodologies, tools, and skills become increasingly central to work in the humanities, questions regarding fundamentals, project outcomes, assessment, and design have become urgent. The specifications provide a set of checklists to guide those who do work in the digital humanities, as well as those who are asked to assess and fund digital humanities scholars, projects, and initiatives.

Keywords: Digital humanities. Methodologies. Projects. Assessment.

O que são as Humanidades Digitais? As Humanidades Digitais se referem a novos modos de produção acadêmica e de unidades institucionais para a pesquisa, ensino e publicações colaborativas, transdisciplinares e permeadas pelas tecnologias computacionais. Elas são menos um campo unificado do que um feixe de práticas convergentes que exploram um universo no qual a mídia impressa não é mais o meio primário no qual o conhecimento é produzido e disseminado.

Os meios, ferramentas e técnicas digitais têm expandido os conceitos tradicionais de conhecimento nas artes, nas humanidades e nas ciências sociais, mas as Humanidades Digitais não tratam somente “sobre” o digital (no sentido de limitar seu escopo ao estudo da cultura digital). Nem as Humanidades Digitais tratam somente “sobre” as humanidades como tradicionalmente compreendidas, uma vez que elas postulam o remapeamento das práticas tradicionais. Em vez disso, as Humanidades Digitais são definidas pelas oportunidades e desafios que surgem da conjunção do termo *digital* como o termo *humanidades* para formar um novo substantivo coletivo.

As oportunidades incluem redesenhar as fronteiras entre as humanidades, as ciências sociais, as artes e as ciências naturais, expandindo o impacto social e o público da pesquisa acadêmica nas humanidades; desenvolvendo novas formas de investigação e produção de conhecimento e revigorando aquelas que haviam sido deixadas de lado; treinando futuras gerações de humanistas através do aprendizado prático e baseado em projetos como um complemento ao aprendizado em sala de aula; e desenvolvendo práticas que expandam o escopo da pesquisa humanística, melhorando sua qualidade e aumentando sua visibilidade.

Esses desafios incluem abordar questões fundamentais tais como: Como podem habilidades tradicionalmente usadas nas humanidades ser reformuladas em termos multimidiáticos? Como e por quem serão definidos os contornos da memória histórica e cultural na era digital? Como podem práticas tais como as narrativas digitais coincidir ou divergir das narrativas orais ou impressas? Qual é o lugar da *humanitas* em um mundo interconectado?

O que define as Humanidades Digitais atualmente? A era computacional tem estado presente desde a Segunda Guerra Mundial, mas após o advento da computação pessoal, da *World Wide Web*, da comunicação móvel e das mídias sociais, a revolução digital entrou em uma nova fase, dando origem a uma esfera pública globalizada vastamente expandida e a possibilidades transformadas de criação e disseminação de conhecimento.

Baseando-se na primeira geração de trabalho das humanidades computacionais, a atividade mais recente das Humanidades Digitais busca revitalizar tradições das artes liberais na linguagem eletrônica do século XXI: uma linguagem na qual, destituído de seu tradicional suporte em papel, o texto está crescentemente aliado a imagens fixas e móveis bem como a sons, e cujos suportes têm se tornado cada vez mais móveis, abertos e extensíveis.

E a própria noção de primazia do texto está sendo desafiada. Enquanto as ondas iniciais das humanidades computacionais se concentravam em tudo, desde os estudos de frequência de palavras e análises textuais (sistemas de classificação, marcação, codificação) à edição de hipertextos e a construção de bases de dados textuais. As Humanidades Digitais contemporâneas marcam um movimento para além do privilégio do textual, enfatizando métodos gráficos de produção e de organização de conhecimento, o design como um componente integral da pesquisa, os entrecruzamentos transmídia, e um conceito expandido do *sensorium* do conhecimento humanístico. Elas também são caracterizadas pelo foco intensificado na construção de ferramentas, ambientes e plataformas transferíveis para o trabalho acadêmico colaborativo e por uma ênfase na curadoria como uma característica definitiva da prática acadêmica.

O que as Humanidades Digitais não são? O mero uso das ferramentas digitais para o propósito da pesquisa e comunicação humanística não se qualifica como Humanidades Digitais. Tampouco, como já observado, devem as Humanidades Digitais ser entendidas como o estudo dos artefatos digitais, novas mídias ou cultura contemporânea no lugar de artefatos físicos, antigas mídias ou cultura histórica.

Pelo contrário, as Humanidades Digitais compreendem o seu objeto de estudo como o registro humano total, da pré-história até o presente. Esse é o motivo pelo qual campos tais como as humanidades clássicas e arqueologia têm desempenhado um papel tão importante no desenvolvimento das Humanidades Digitais, como o têm desempenhado, por exemplo, os estudos de mídia. Esse é também o motivo pelo qual alguns dos maiores setores das pesquisas em Humanidades Digitais se estendem para fora do núcleo tradicional das humanidades para adotar métodos quantitativos das ciências sociais e naturais, bem como as técnicas e modos de pensar das artes.

De onde vêm as Humanidades Digitais? As raízes do trabalho computacional nas humanidades remontam a 1949, quando o acadêmico jesuíta Roberto Busa, trabalhando em colaboração com a IBM, empreendeu a criação de uma abordagem automatizada do seu vasto *index thomisticus*, um índice gerado computacionalmente dos escritos de Tomás de Aquino. Por meio desses primeiros usos de mainframes para automatizar tarefas, tais como busca de palavras, ordenação, contagem e listagem, acadêmicos puderam processar *corpora* textuais em uma escala impensável com métodos anteriores, que se baseavam nos cartões de indexação escritos à mão ou datilografados.

Outros projetos iniciais incluíram a estreia, em 1966, do *Computers and the Humanities*, o primeiro periódico acadêmico especializado na área. Sete anos depois, a Associação para a Computação Literária e Linguística (ALLC) foi fundada, seguida pela Associação para Computadores e as Humanidades (ACH) em 1978. Em meados dos anos 1980, métodos computacionais para análise linguística tinham se tornado suficientemente difundidos a ponto de serem necessários protocolos para taguear textos digitais. Isso estimulou o desenvolvimento da *Text Encoding Initiative* (TEI). Essa importante empreitada reformulou o campo de pesquisas acadêmicas sobre textos eletrônicos e levou a subsequente edição digital a ser executada em *Extensible Markup Language* (XML), o esquema de taguamento do qual a TEI é um subconjunto especializado.

Os primeiros experimentos em humanidades com estruturas de bancos de dados e edição hipertextual estruturada em links e nós (ao invés das convenções lineares da mídia impressa) datam desse período, bem como os muitos projetos piloto das humanidades computacionais nos Estados Unidos, patrocinados pela *National Endowment for the Humanities* e outras agências, organizações e fundações.

Como a web e outras redes afetam as Humanidades Digitais? Enquanto essa revolução nos protocolos acontecia, a explosão da computação pessoal em meados da década de 1980, combinada com o advento da *World Wide Web* uma década mais tarde, deram surgimento a uma nova geração de trabalho nas Humanidades Digitais, que era menos centrado em textos e mais direcionado ao design. O ambiente do desktop – com sua interface gráfica de usuário, ferramentas de *wysiwyg*⁸ em tempo real e a evolução das linhas de comando para ícones e interfaces baseadas em janelas – não somente expandiu grandemente o *corpus* dos documentos originalmente digitais, mas também inaugurou a integração gradual com o áudio, o vídeo e os gráficos.

⁸ Nota da Tradutora. A sigla *wysiwyg* é o acrônimo da expressão idiomática *What You See Is What You Get*, ou seja, *O que você vê é o que você obtém*, que foi adotada no âmbito computacional quando os processadores de texto passaram a apresentá-lo na tela em sua forma final, ou seja, indicando que o que estava sendo mostrado correspondia ao produto final.

Essa integração amadureceu durante as últimas décadas e deu à cultura Web seu profundo caráter multimídia. Ela também favoreceu o aprimoramento de modelos de compartilhamento, cocriação, publicação e construção comunitária que tem situado a Web no centro dos debates sociais contemporâneos e dos processos socioeconômicos. O conceito de Web, como uma esfera pública que estende os espaços públicos físicos da vida contemporânea tem, é claro, sido intensificado graças a smartphones, tablets e outros aparatos ubíquos e pervasivos de mídia e computação.

O que está à frente das Humanidades Digitais? As Humanidades Digitais contemporâneas não se opõem ao passado, mas se fundamentam nele. Elas honram os trabalhos pioneiros levados a cabo durante as sete últimas décadas na forma de processamento estatístico (linguística computacional), *linking* (hipertexto), modelagem (arquiteturas e exibição visual), a criação de dados estruturados (XML) e a edição iterativa e controle de versões (para edições críticas bem como para análises e práticas criativas), mesmo quando ele busca ir além da construção de repositórios e edição para novas práticas sintéticas. Ele é inspirado pela mesma convicção fundamental que animava as humanidades computacionais e os pioneiros dos primórdios das Humanidades Digitais: a convicção de que as ferramentas computacionais têm o potencial de transformar o conteúdo, o escopo, as metodologias e a audiência da investigação humanística.

Questão 2: o projeto como unidade básica

Por que projetos? Projetos são tanto substantivos como verbos: Um projeto é um tipo de produção acadêmica que requer design, administração, negociação e colaboração. Também é trabalho acadêmico que projeta, no senso de futuridade, como algo que ainda não é. Projetos são geralmente realizados em times, com colaboradores trazendo conjuntos de habilidades e interesses complementares a fim de conceitualizar as questões de pesquisa que estão sendo investigadas e projetar possíveis trajetórias para respondê-las. Portanto, projetos são projetivos, envolvendo processos iterativos e muitas dimensões de coordenação, experimentação e produção.

Quem está envolvido nos projetos de Humanidades Digitais? Projetos de Humanidades Digitais envolvem tipicamente múltiplos círculos de pesquisadores, da faculdade e administração a estudantes e parceiros da comunidade. A complexidade e a escala de um projeto geralmente implicam o envolvimento de múltiplos estratos de pessoal de dentro e através de instituições de ensino.

Projetos podem envolver instituições parceiras, tais como museus, bibliotecas, e arquivos, bem como membros da comunidade de alunos, e membros de redes virtuais interessadas, tais como colecionadores, historiadores amadores e afins. Parcerias com corporações, em particular companhias de mídia e tecnologia, também são possíveis, embora as culturas corporativa e acadêmica possam ser diferentes em seus objetivos e valores.

Como são organizados os projetos de Humanidades Digitais? Projetos são usualmente iniciados pelas faculdades, pelas equipes ou pelos estudantes. Eles são geralmente construídos em torno de uma questão de pesquisa e/ou uma coleção ou repositório de arquivos da universidade. Muitos acontecem fora da sala de aula; outros envolvem um projeto de pesquisa ancorado em um curso recorrente.

Um Investigador Principal (ou IP), co-IPs, conselheiros do projeto, equipes, estagiários e estudantes são todos parte do time do projeto. É responsabilidade do IP organizar o time do projeto, estabelecer datas para entregas, e avaliar o projeto a cada estágio de desenvolvimento.

Qual é a diferença entre projetos de Humanidades Digitais e projetos das Grandes Humanidades? Projetos de Humanidades Digitais podem ter todos os tamanhos: grandes, médios e pequenos. Alguns dos primeiros e definidores projetos de Humanidades Digitais, contudo, bem como trabalhos contemporâneos proeminentes, têm assumido a forma de projetos de Grandes Humanidades, que são realizados ao longo de muitos anos, com muitos colaboradores, desenvolvedores e patrocinadores envolvidos nos vários estágios do desenvolvimento. Projetos de Grandes Humanidades são construídos ao longo das linhas da Grande Ciência. Eles envolvem iniciativas baseadas em equipes de larga escala e longo prazo, que constroem grandes perspectivas a partir de blocos de conhecimento especializado. Os pesquisadores e os membros do time, de historiadores a tecnólogos e designers, podem chegar às centenas.

Pequenos ou diminutos projetos de Humanidades Digitais são tipicamente levados à frente por indivíduos ou pequenos times em consultoria com equipes experientes. À medida que as plataformas e protocolos padrão têm emergido, editar, construir a forma de exibição, a análise das redes e o desenvolvimento dos repositórios requerem menos investimentos pontuais.

A grande maioria dos projetos em Humanidades Digitais se situa entre as duas pontas do espectro.

Como as Humanidades Digitais dão continuidade às formas tradicionais de pesquisa e ensino nas humanidades? Assim como as pesquisas e o ensino baseados nas humanidades tradicionais, o trabalho nas Humanidades

Digitais envolve práticas de análise, crítica e interpretação; editoração e anotação; pesquisa histórica e contextualização. Ele examina as propriedades formais e históricas dos trabalhos da imaginação, da interação entre o sujeito e a sociedade, a história das ideias e a cultura material. Ele trata de características qualitativas e não quantificáveis da experiência humana: complexidade, ambiguidade, especificidade dos meios e subjetividade. Ele se baseia em abordagens tradicionais de estudo, preservação e classificação de *corpora* culturais.

Embora a variedade de mídias com as quais as Humanidades Digitais trabalham se estenda para além da textual, seus comprometimentos principais se harmonizam com os duradouros valores da tradição humanística: a busca de acuidade e de clareza analítica, a construção de argumentos efetivos, o uso rigoroso da evidência e a expressividade e eficácia comunicativa.

As Humanidades Digitais trabalham com conjuntos de dados grandemente expandidos, através das mídias e por meio de novos casamentos entre o digital e o físico, resultando em definições de e compromisso com o conhecimento que abrangem todo o *sensorium* humano.

Tanto o estudo tradicional em sala de aula quanto o individual permanecem aspectos chave da paisagem do aprendizado em Humanidades Digitais. Ao mesmo tempo, muitos precedentes para o trabalho colaborativo nas comunidades de letras e redes de conhecimento são aumentados pelas plataformas digitais em um tecido animado por oportunidades de aprendizado prático e baseado em projetos.

Desde a antiguidade, os modelos dominantes de investigação humanística têm favorecido o entendimento do trabalho intelectual como solitário e contemplativo, apartado – e até mesmo superior – ao trabalho manual e ao campo do fazer ou produzir. As Humanidades Digitais reinserem esses modelos em um modelo aumentado de pedagogia que enfatiza o aprendizado através do fazer e produzir, seja no nível individual ou de grupo.

Em que as Humanidades Digitais são descontínuas em relação às formas tradicionais de pesquisa e ensino nas humanidades? Por aproximadamente seis séculos os modelos humanísticos de conhecimento têm sido formatados pelo poder da impressão gráfica como o meio primordial para a produção e disseminação do conhecimento. Ao invés de rejeitar a cultura impressa ou abraçar a simples adoção dos modelos de impressão em moldes digitais, as Humanidades Digitais estão engajadas em desenvolver modelos de conhecimento impressão-mais [*print-plus*] e pós-impres-

são *[post-print]*. Ambos envolvem mais que uma atualização do sistema de entrega do conhecimento. Eles implicam a reformulação cognitiva e epistemológica dos campos humanísticos como uma função dos recursos disponibilizados pelo digital em relação ao impresso. Eles também dizem respeito ao aumento do papel do trabalho de equipe e jogo colaborativo na pesquisa e treinamento em humanidades.

Como as Humanidades Digitais funcionam na era impressão-mais [print-plus]? A impressão tipicamente oferece um único ângulo de visão, a organização linear, um resultado investigativo caracterizado pela finitude e estabilidade, e uma escala de documentação e argumentação que precisa respeitar as proporções físicas do livro. A era digital impressão-mais *[print-plus]*, em contraste, permite alternar idas e vindas entre múltiplas visões dos mesmos materiais. Ela permite mudanças fluídas de escala, para possibilitar o zoom do macro ao micro nível e do entrelaçamento de conjuntos de dados (tais como materiais fonte, notas e correspondências) para resultados investigativos. As telas e espaços aumentados da era impressão-mais *[print-plus]* permitem facetar, filtrar e dar diferentes versões dos *corpora*; a coexistência de múltiplos caminhos dentro de um mesmo repositório; e formas multilineares de argumento. Ela é extensível no duplo sentido de permitir escalas aparentemente ilimitadas e de processos de entitativos *[being process]* – ao invés de baseado em produtos. Quando um livro vai para impressão, ele se estabiliza em uma edição que precisa ser reeditada a fim de ser revisada; um artefato pode ser alterado ou revisado em um substrato reescritível, que suporta rápidas quotas de atualização. O mesmo artefato digital pode ter múltiplas vidas em múltiplas plataformas, com múltiplos autores. Ele pode ser remixado por outros, antes, durante e depois de sua conclusão.

Como são os projetos de Humanidades Digitais financiados e sustentados? Porque eles cruzam fronteiras entre disciplinas, entre conhecimento teórico e aplicado, e entre as humanidades, biblioteconomia, tecnologias da informação e design, os projetos de Humanidades Digitais tipicamente requerem estruturas de suporte que atravessam os departamentos e linhas organizacionais acadêmicas convencionais. Fundações privadas, agências públicas de fomento e indústrias parceiras têm fornecido fundos para projetos em todas as escalas.

Financiamento para a pesquisa nas humanidades é muito mais limitado do que nos campos da ciência e engenharia, mas o escopo e caráter inovador das Humanidades Digitais têm conduzido muitos projetos a angariar financiamentos externos com sucesso. A fim de atrair e susten-

tar tais financiamentos, provou-se essencial para os projetos receber apoio interno durante o período de incubação a fim de que eles possam provar seu valor ao atingir com sucesso um conjunto inicial de referências.

Sustentar tais projetos requer que a universidade e os alunos que assumem a posição de liderança precisam o apoio e o reconhecimento de que esse trabalho é uma combinação de pesquisa, ensino e prestação de serviços.

Quais são as convenções prevalentes de creditação e atribuição e de modelos de autoria para os projetos de Humanidades Digitais? Práticas de creditação e autoria tradicionais nas humanidades são baseadas na autoria única. Embora práticas de atribuição ainda sejam fluídas na comunidade das Humanidades Digitais, o modelo emergente reconhece que muitos, se não a maioria, dos projetos em Humanidades Digitais são análogos tanto aos projetos de laboratório em ciências naturais, como ao sistema de atribuição colaborativa utilizado nas artes performáticas. Nenhum sistema de creditação padronizado para os projetos em Humanidades Digitais foi adotado universalmente. Mas a tendência dominante vai na direção da diferenciação de papéis, tais como investigador principal, pesquisador, designer, programador, modelador, editor e afins.

Instituições e pragmática

Como os projetos em Humanidades Digitais interconectam a sala de aula com bibliotecas, museus e arquivos? A maioria das faculdades e universidades têm extensos recursos para a pesquisa e estudo que são subutilizados após atingir suas principais missões de pesquisa, ensino e treinamento. As Humanidades Digitais contemporâneas canalizam essas riquezas ao expandir o conceito de sala de aula para abranger bibliotecas, museus e coleções de arquivos, posicionando-as como lugares centrais de treinamento através de pesquisa prática na companhia de pares.

Tanto como nos laboratórios de ciência natural, estudantes envolvidos em projetos de Humanidades Digitais aprendem fazendo e produzindo, trabalhando com essa sala de aula estendida sob a orientação de curadores, arquivistas e pesquisadores experientes, e na companhia de pares. Ao passo que os modelos tradicionais de treinamento humanístico veem a aquisição de conjuntos de habilidades e treinamento em disciplinas como pré-condições para a transição de se engajar na criação de produção acadêmica original, o trabalho em Humanidades Digitais acelera esse aprendizado, inserindo os estudantes em comunidades de pesquisa desde o início.

Como podem os projetos de Humanidades Digitais envolver a colaboração interuniversitária? A escala e o escopo de muitos projetos de Humanidades Digitais, bem como suas ligações a coleções físicas e necessidades de infraestrutura em TI, os tornam idealmente adequados para a colaboração interuniversitária. Projetos podem ser desenvolvidos e divididos estrategicamente entre múltiplas instituições parceiras, alavancando forças específicas, distribuindo cargas de trabalho, partilhando os benefícios dos resultados da pesquisa, e construindo pontes interinstitucionais.

Os benefícios incluem a divisão de custos e perspectivas aprimoradas de financiamento externo. Eles também transcendem a esfera prática: eles permitem modelos de pesquisa de Grandes Humanidades cujos resultados são de potencial interesse para amplas audiências interdisciplinares e não especialistas. Ao envolver múltiplas instituições, tais projetos contribuem para um senso de identidade compartilhada e de pertencimento a uma comunidade de pesquisa mais ampla. Eles também ajudam a responder ansiedades estudantis específicas em relação ao valor prático do conhecimento e pesquisa em humanidades.

Como podem os projetos em Humanidades Digitais envolver expertise fora da academia? Muitos projetos de Humanidades Digitais se desenvolvem inteiramente dentro de uma única faculdade ou universidade. Mas outros requerem domínios de conhecimento e formas de expertise que estão sub-representados ou que permanecem fora dos limites do campo acadêmico. Combinar expertise intra e extramuros dentro de projetos em Humanidades Digitais bem planejados é sempre essencial ao seu sucesso. Tais abordagens incluem trabalho com comunidades de colecionadores e associações históricas, e o uso de vastos recursos humanos para o processamento, transcrição e anotação dos documentos arquivados. Não somente pode o escopo e qualidade da pesquisa em humanidades se beneficiarem de tais parcerias, como também eles contribuem para a criação de novas classes de cidadãos acadêmicos, que de outra forma seriam meros cidadãos consumidores.

Como podem as parcerias extramuros desempenhar um papel em desenvolver, apoiar e sustentar projetos em Humanidades Digitais? A promoção do conhecimento público é um valor essencial às Humanidades Digitais. Parcerias extramuros – seja com sociedades profissionais, associações históricas, instituições de aprendizado informal (bibliotecas, museus, arquivos), corporações ou entidades públicas – podem estender o alcance e o impacto na pesquisa em humanidades na sociedade contemporânea. As parcerias mais bem sucedidas se dirigem a questões de interesse crí-

tico compartilhado com resultados de pesquisa que se elevam aos mais altos padrões de rigor acadêmico, ao mesmo tempo em que são conjugados através de plataformas em múltiplas mídias na “linguagem” das instituições parceiras por meio de exposições, performances, livros, publicações web, e outros meios.

Parcerias podem expandir a profundidade e diversidade do conjunto de talentos de participantes em um projeto, ampliando o público potencial do projeto e seu impacto e, como na colaboração interuniversitária, ajudar a solidificar a sustentabilidade financeira em curto e longo prazo.

Como podem as instituições educacionais dar apoio à pesquisa em Humanidades Digitais? Os projetos de pesquisa em Humanidades Digitais requerem fronteiras fluídas em departamentos acadêmicos e unidades institucionais. Porque os projetos são sempre baseados em times e implicam modelos mistos de conhecimento teórico e aplicado através de domínios tradicionalmente separados de “pesquisa”, “ensino”, e “serviço”, tais elementos como facilidades de design, sistemas de informação, produção multimídia, trabalho em TI, e pesquisa baseada em coleções não são meros “apoios”, mas sim características integrais do planejamento e execução de projeto.

Além de promover uma cultura onde tais fronteiras não obstaculizem a inovação, as instituições devem adotar o coensino como um padrão da nova paisagem nas humanidades, ao invés de penalizá-lo como uma forma de redução de trabalho. A cocriação precisa ser vista como uma forma legítima de trabalho intelectual de alunos e acadêmicos, complementar às formas tradicionais de resultado. A facilidade de acesso e as restrições de uso em museus, arquivos, e coleções especiais de bibliotecas representam uma pré-condição chave para a criação de uma sala de aula expandida e prática, e servem para que modelos de acesso aberto equacionem a preservação com a proliferação, ao invés de controle restrito.

Departamentos jurídicos de universidades e faculdades devem ser cuidadosos para não interpretar restrições de direitos autorais de forma restrita por falta de vontade de intermediar riscos hipotéticos. O uso justo precisa ser entendido no sentido mais amplo possível a fim de não algemar a pesquisa em Humanidades Digitais. Coleções alocadas em faculdades e universidades precisam ser compartilhadas com a comunidade de pesquisadores tão gratuitamente quanto possível. Por último, mas não menos importante, instituições de ensino superior devem promover e nutrir uma cultura menos avessa a riscos nas disciplinas das humanidades: uma cultura onde, como nas ciências, “falhas” sejam aceitas como

resultado produtivo ao empreender trabalho inovador e especulativo. Diferenciar entre falhas produtivas e pesquisa pouco consistente são essenciais para promover comunidades de pesquisa onde a inovação é um valor fundamental.

Quais são os nichos institucionais que dão melhor suporte a projetos em Humanidades Digitais? Projetos em Humanidades Digitais geralmente florescem menos dentro de departamentos únicos, escolas ou unidades institucionais do que através de tais estruturas. Mesmo centros de pesquisa em humanidades construídos para abrigar e dar apoio à pesquisa individual nem sempre provaram ser ideais – embora alguns tenham reformulado suas políticas, modelos de financiamento e infraestrutura física com sucesso para dar suporte a trabalhos colaborativos em Humanidades Digitais.

Mais tipicamente as Humanidades Digitais têm prosperado em laboratórios ou centros independentes e livres, onde existe uma comunidade de acadêmicos (humanistas e não humanistas), membros de equipe, curadores e estudantes interessados na exploração compartilhada de modelos inovadores de produção acadêmica. Tais ambientes são mais bem vistos como espaços híbridos de fazer, pensar e entreter combinando facilidades computacionais, imagens digitais, facilidades de produção de som e vídeo e espaços de encontro e exposições.

Como podem as instituições avaliar a escala de investimento e expectativas para projetos de Humanidades Digitais apropriadas a elas? Projetos têm todos os tamanhos. Não há nenhuma razão inerente pela qual um grande projeto não possa ser empreendido por uma instituição pequena ou um pequeno projeto por uma grande instituição. Tampouco há nenhuma razão inerente pela qual acadêmicos individuais não possam assumir colaborações de larga escala entre múltiplas faculdades ou universidades.

Não há nenhuma fórmula para o sucesso. A escala e a forma que os projetos em Humanidades Digitais tomam devem ser ditadas pelo planejamento cuidadoso do projeto – combinando questões, ambições e resultados investigativos antecipados – bem como pelos recursos financeiros, pessoais e logísticos disponíveis. Tanto quanto nas ciências laboratoriais, isso implica em um equilíbrio entre uma visão pragmática e uma iniciativa empreendedora.

Como podem líderes e pares acadêmicos avaliar projetos em Humanidades Digitais? Métricas para avaliar a qualidade e o impacto de projetos em Humanidades Digitais combinam métodos de avaliação tradicionais nas humanidades com novos fatores. A revisão por pares permanece fundamental para o processo de avaliação, mas agora se baseiam tanto

na comunidade de praticantes líderes em Humanidades Digitais, bem como em pares da mesma área. Uma cultura menos avessa a riscos é o pré-requisito para que um modelo dirigido a uma maior inovação – e experimentação – nas Humanidades Digitais possa ganhar espaço.

Além dos critérios tradicionais baseados em pares, algumas ferramentas de avaliação que têm uma longa história nas ciências naturais e sociais podem se tornar relevantes aos campos das humanidades: citações, sucesso ao receber financiamentos, impactos públicos e afins. Deve ser observado que variações no tamanho dos campos tornam a cautela essencial no uso de ferramentas quantitativas; do contrário elas proveirão medidas muito cruas e possivelmente enganosas da importância ou impacto. Produção acadêmica original e rigor intelectual permanecem a essência do trabalho em Humanidades Digitais.

Métricas tradicionais de produtividade baseadas em impressão já estão sendo eclipsadas pelas realidades da publicação digital e impressão-mais [*print-plus*], de forma que as expectativas de produtividade devem abranger múltiplas mídias, diferentes formatos e variadas escalas de contribuição ao conhecimento. Em outras palavras, as mídias e tecnologias nas quais o trabalho intelectual é realizado importam tanto quanto seu “conteúdo”. Isso significa que o “trabalho” não é somente o conteúdo, mas, ao invés disso, *tudo*: o ambiente que foi planejado para a performance do trabalho e sua publicação, a interface e a estrutura de dados, o banco de dados interno e o código que permite múltiplas formas de engajamento por parte da audiência. Tudo isso importa em avaliações de qualidade e rigor.

Especificação 1: como avaliar a produção acadêmica digital

O seguinte fornece um conjunto de coordenadas para a avaliação de produção acadêmica digital nas humanidades, ciências sociais, artes e disciplinas relacionadas. As coordenadas são dirigidas, principalmente, a comitês de revisão acadêmica, catedráticos, decanos e reitores que querem saber como avaliar a produção acadêmica digital.

Esta lista também pretende dar informações para o desenvolvimento de políticas institucionais amplas para apoiar e avaliar a produção acadêmica e o trabalho criativo que reflete valores tradicionais, ao passo que também incorpora entendimentos específicos de novas plataformas e formatos.

Fundamentos para uma avaliação inicial. O trabalho deve ser avaliado no meio no qual ele foi produzido e publicado. Se ele é um website, isso significa vê-lo em um browser com os plug-ins apropriados necessários para o site funcionar. Se ele for um modelo de simulação virtual, isso pode significar ir ao laboratório equipado com os softwares necessários e sistemas de projeção para visualizar o modelo. Trabalho que é baseado em tempo – tais como vídeos – serão geralmente representados por imagens fixas, mas os avaliadores também precisam dar atenção a cliques a fim de avaliar completamente o trabalho. O mesmo pode ser dito para a interface de desenvolvimento, uma vez que imagens fixas não podem demonstrar completamente a natureza interativa da interface de pesquisa. Autores de trabalhos digitais devem prover uma lista de requisitos do sistema (tanto de hardware quanto de software, incluindo browsers, versões e plug-ins compatíveis) para a visualização do trabalho. É incumbência dos departamentos acadêmicos de pessoal verificar que as tecnologias estejam disponíveis e instaladas nos sistemas que serão utilizados pelos avaliadores antes que eles avaliem o trabalho digital.

Creditação. Projetos digitais são geralmente de natureza colaborativa, envolvendo times de acadêmicos que trabalham juntos em diferentes locais em variados períodos de tempo. Autores de trabalhos digitais devem prover uma articulação clara do papel ou papéis que eles desempenharam na gênese, desenvolvimento e execução de projeto digital. É impraticável – se não impossível – separar cada micro contribuição feita pelos membros do time uma vez que projetos digitais são geralmente sinérgicos, iterativos, experimentais e até mesmo dinamicamente gerados através de colaborações contínuas. Entretanto, os autores devem indicar os papéis que eles desempenharam (e períodos de tempo) em cada fase do desenvolvimento do projeto. Quem idealizou o projeto e planejou as especificações iniciais (funcionais e técnicas)? Quem criou os modelos? Quem escreveu as propostas de financiamento ou assegurou o financiamento que deu suporte ao projeto? Qual papel cada colaborador desempenhou no desenvolvimento e execução do projeto? Quem foi o autor do conteúdo? Quem decidiu como o conteúdo seria acessado, exibido e armazenado? Qual é a “face pública” do projeto e quem a representa e como?

Rigor intelectual. Projetos digitais variam grandemente e podem não “parecer” produção acadêmica tradicional; ao mesmo tempo, o rigor acadêmico deve ser avaliado, examinando-se como o trabalho contribui para e avança o estado do conhecimento em um dado campo ou campos. Qual é a natureza do novo conhecimento criado? Qual é a metodologia usada

para criar esse novo conhecimento? É importante para comitês avaliadores reconhecer que conhecimento novo não é somente conteúdo novo, mas também novas formas de organizar, classificar e interagir com o conteúdo. Isso significa que parte da contribuição intelectual de um projeto digital é o desenho de sua interface, a base de dados e o código, os quais governam a forma do conteúdo. Acadêmicos digitais não estão somente na posição de realizar pesquisas originais, mas também de inventar novas plataformas acadêmicas. Quinhentos anos de impressão naturalizaram tão completamente a “aparência” do conhecimento que pode ser difícil para o avaliador compreender completamente essas novas formas de documentação e o esforço intelectual necessário a seu desenvolvimento. Esse é o duplo fardo – e a dupla oportunidade – para a criatividade no domínio digital.

Combinando pesquisa, ensino e serviço. Projetos digitais quase sempre têm múltiplas aplicações e usos que melhoram a pesquisa, o ensino e a prestação de serviços. Projetos digitais de pesquisa podem dar contribuições transformadoras para a sala de aula e, às vezes, até terem um impacto no grande público. Esse efeito cascata não deve ser diminuído. Comitês de avaliação necessitam atentar para colegas que diminuem as contribuições investigativas do trabalho digital ao cavalheiramente caracterizá-lo como mera “ferramenta” para o ensino ou serviço. Ferramentas formam conhecimento e conhecimento formata ferramentas. Mas também é importante que comitês de avaliação foquem nas contribuições para a pesquisa do trabalho digital fazendo perguntas tais como: Como o trabalho se engaja com um problema específico de uma disciplina acadêmica ou grupo de disciplinas? Como o trabalho reformula aquele problema ou contribui para uma nova maneira de entendê-lo? Como o trabalho avança um argumento através tanto do conteúdo como da forma em que o conteúdo é apresentado? Como o planejamento da plataforma constitui um argumento? Para responder a esta última questão, comitês avaliadores devem pedir documentação descrevendo o processo de desenvolvimento e o planejamento da plataforma ou software, tais como esquemas de bases de dados, desenhos de interface, módulos de codificação (e explicação do que eles fazem), bem como amostras de tipos de dados. Se o projeto é de fato primariamente destinado ao ensino, como ele transformou o ambiente de aprendizado? Quais contribuições ele deu para o ensino e como têm essas contribuições sido avaliadas?

Avaliação por pares. Projetos digitais devem ser avaliados por pares acadêmicos em áreas que estão aptas a avaliar a contribuição do projeto para o conhecimento e a situá-lo em uma paisagem intelectual relevante. A avaliação por pares pode ocorrer formalmente através de cartas de solicitação, mas pode também ser avaliada através de fóruns online, citações e

discussões em locais acadêmicos, por patrocínios recebidos de fundações e outras fontes de financiamento, e através de apresentações públicas do projeto em conferências e simpósios. O projeto gerou publicações em periódicos avaliados por pares ou ganhou prêmios de associações profissionais? Como ele se compara a projetos equivalentes no mesmo campo que usam ou desenvolvem tecnologias similares ou tipos similares de dados? Finalmente, os financiamentos recebidos são geralmente indicadores significativos de avaliação por pares. É importante que os avaliadores se familiarizem com as organizações de financiamento através de escolas e disciplinas, incluindo as humanidades, as ciências sociais, as artes, estudos informacionais e biblioteconomia e as ciências naturais, uma vez que estes são indicadores de prestígio e impacto.

Impacto. Projetos digitais podem ter um impacto em numerosos campos da academia, bem como em instituições e até no público em geral. Eles geralmente cruzam o limite que se ergue entre pesquisa, ensino e serviço de formas inovadoras. O impacto pode ser medido de muitas maneiras, incluindo as seguintes: apoio de fundações ou agências de financiamento, número de visualizadores e colaboradores de um site e o teor das contribuições, citações tanto na literatura tradicional como online (blogs, mídia social, links e *trackbacks*), uso ou adoção do projeto por outros acadêmicos ou instituições, conferências e simpósios acerca do projeto, e ressonância pública e alcance comunitário (tais como exposições em museus, impacto em políticas públicas, adoções em currículos, e assim por diante).

Aproximando equivalências. Um projeto de pesquisa digital é “equivalente” a um livro publicado por uma editora universitária, um volume organizado ou um artigo de pesquisa? Esses tipos de questões são geralmente enganadores uma vez que eles se fundamentam na comparação entre artefatos fundamentalmente diferentes de conhecimento e, talvez mais problematicamente, consideram publicações impressas como a norma e a referência a partir das quais medir todos os outros trabalhos. Revisores devem ser capazes de avaliar a significância do trabalho digital baseado em uma série de fatores: a qualidade e a quantidade de pesquisa que contribuiu para o projeto; a extensão de tempo gasto e o tipo de investimento intelectual dos criadores e colaboradores; o alcance, profundidade e formatos de tipos de conteúdo e as maneiras pelas quais esse conteúdo é apresentado; e a natureza da autoria e do processo de publicação. Projetos de larga escala com grandes financiamentos, múltiplos colaboradores e um amplo espectro de resultados acadêmicos podem justificadamente receber mais peso no processo de avaliação e promoção do que projetos de pequena escala e curto prazo.

Ciclos de desenvolvimento, sustentabilidade e ética. É importante que os comitês avaliadores reconheçam a natureza iterativa dos projetos digitais, que podem implicar múltiplas revisões em diversos ciclos de revisão, à medida que os projetos crescem, mudam e amadurecem. Dado que os ciclos de avaliação acadêmica são geralmente separados por vários anos (enquanto os avanços digitais ocorrem mais rapidamente), avaliadores devem considerar projetos individuais em seus contextos específicos. Em qual “estágio” está o projeto em sua forma atual? É ele considerado “completo” pelos criadores ou ele vai continuar em novas iterações, talvez por meio de subprojetos e desenvolvimentos adicionais? Tem o projeto seguido as melhores práticas à medida que elas têm sido estabelecidas no campo, em termos de coleta de dados e produção de conteúdo, uso de padrões e documentação apropriada? Como vai o projeto “viver” e ser acessível no futuro, e qual tipo de infraestrutura será necessário para seu suporte? Aqui, necessidades específicas do projeto e obrigações institucionais se aliam nos mais altos níveis e devem ser discutidas abertamente com decanos e diretores, equipes de bibliotecas e de TI, e líderes do projeto. Finalmente, os projetos digitais podem levantar questões críticas e éticas sobre a natureza e o valor da preservação cultural, história pública, cultura participativa e acessibilidade, diversidade digital e curadoria de coleções que devem ser cuidadosamente consideradas pelos líderes do projeto e comitês avaliadores.

Experimentação e tomada de riscos. Projetos digitais nas humanidades, ciências sociais e artes compartilham, com as práticas experimentais nas ciências, o desejo de serem abertos em relação à iteração e a resultados negativos. Como tal, experimentação e tentativa e erro são partes inerentes da pesquisa digital e devem ser reconhecidas. Os processos de experimentação devem ser documentados e podem provar ser essenciais aos processos de desenvolvimento de uma ideia ou projeto em longo prazo. Propostas [*white papers*], conjuntos de melhores práticas, novos ambientes de design e publicações podem resultar de tais projetos e estes devem ser considerados no processo de avaliação. Experimentação e tomada de riscos na produção acadêmica representam o melhor que a universidade, em todas suas muitas disciplinas, tem a oferecer para a sociedade. Tratar a produção acadêmica que aceita riscos e o desafio da experimentação como uma atividade de valor secundário (ou nulo) para o fomento e o avanço podem servir somente para reduzir a inovação, recompensar a mediocridade e retardar o desenvolvimento da pesquisa.

Especificação 2: produção acadêmica baseada em projetos

A produção acadêmica baseada em projetos exemplifica os princípios das Humanidades Digitais contemporâneas. Ela difere da publicação acadêmica tradicional ao ser baseada em times, distribuída em sua produção e resultados, dependente de recursos trabalhados em rede (técnicos e/ou administrativos) e por ser iterativa e contínua, ao invés de fixa ou final em seus resultados. Ela necessariamente envolve muitas dimensões da concepção, planejamento, coordenação e uso de recursos que constroem camadas extra de complexidade na abordagem tradicional da pesquisa em humanidades. A lista a seguir é útil para a criação de uma proposta de financiamento ou um plano de pesquisa para o trabalho baseado em projetos e reflete os padrões das melhores práticas (com a ressalva de que o debate continua).

Contribuição ao conhecimento. O projeto deve atender os critérios de qualquer trabalho acadêmico através de sua contribuição para o conhecimento em uma disciplina ou campo. Como o projeto dialoga com uma questão ou tópico em um dado campo disciplinar e como ele move o discurso adiante de uma forma inovadora? O projeto contribui para e avança o estado de conhecimento de um dado campo ou campos?

O modelo de conhecimento. Como é modelado e formatado o conhecimento: como um argumento, uma apresentação, uma exibição? O que pode ser considerado no projeto como um princípio teórico, método ou informação que é útil a outros acadêmicos, incluindo aqueles que não estão engajados com a pesquisa em Humanidades Digitais? Como o projeto modela e corporifica novos conhecimentos?

Questões investigativas e mídia digital. Ambientes digitais permitem diferentes abordagens para relacionar e processar materiais e isso deve ser demonstrado no plano de pesquisa. Simplesmente colocar algo online não é pesquisa digital. O teste decisivo é perguntar o que está sendo feito que não poderia ser realizado em um trabalho acadêmico tradicional e baseado em impressão gráfica. Como o projeto de pesquisa foi formulado a partir de recursos de métodos digitais?

Ferramentas e conteúdo. Muitos projetos digitais envolvem a recombinação e reconfiguração inovadora de ferramentas existentes para a formulação de novo conhecimento. É esse um projeto baseado em ferramentas ou um projeto orientado pelo conteúdo, e como estes se intersectam? Como pode o trabalho intelectual de planejamento e desenvolvimento de uma “ferramenta” ser avaliado em conjunto com o “conteúdo”? Em qual medida eles são inseparáveis e por que?

Métodos. O projeto tem uma tese ou princípio metodológico orientador? Como a plataforma digital permitiu explorá-lo, testá-lo, defendê-lo, demonstrá-lo ou mesmo refutá-lo?

Artefato produzido digitalmente ou digitalizado. Projetos digitais geralmente combinam materiais analógicos que foram escaneados ou digitalizados e elementos originariamente digitais – análises, pesquisa, processamento, ou arquivos de nova autoria. Elementos de estrutura de informação são também produzidos digitalmente. Como são cada um desses elementos entendidos e qual papel eles desempenham no projeto como um todo?

Licenciamento e compartilhamento de coleções. O futuro do aprendizado humanístico e o nível de impacto social que a produção acadêmica em humanidades pode atingir dependem do acesso irrestrito a repositórios culturais e históricos; por conseguinte, as licenças menos restritivas devem ser a norma. Que tipos de questões sobre licenciamento e propriedade intelectual o projeto encontrará? Como pode o trabalho ser acessado e utilizado pela comunidade acadêmica e pelo grande público?

Interface como representação de conhecimento e modelagem de conteúdo. A interface de um projeto expressa um argumento em seu design. Ela oferece uma sinopse dos conteúdos do projeto, ou um conjunto de pontos de entrada para atividades que podem ser realizadas? Compreender as maneiras como a interface é estruturada, como ela corporifica as ideias do projeto, e como ela dá suporte ao engajamento com o projeto é essencial.

Gestão de equipe, colaborativa e de projeto. Saber quem será responsável por cada parte de um projeto digital é crucial para o desenvolvimento e design. O papel de cada participante deve ser especificado na documentação: a concepção do projeto, o plano de pesquisa, a análise técnica, o desenvolvimento web (infraestrutura), o design web (interface), o desenvolvimento de conteúdo, o design de base de dados, e assim por diante. Algum cálculo da porcentagem de esforço no projeto como um todo deve ser indicada.

Créditos de contribuição intelectual e autoria. Equipes de projeto devem trabalhar colaborativamente e a atividade de pesquisa se desenvolve com a implementação; ela não é separada dela. Mas a responsabilidade pela questão de pesquisa e pela contribuição intelectual de cada participante deve ser claramente apresentada na documentação. Isso deve incluir a descrição de como o projeto foi formatado por decisões de design, conhecimento específico de disciplinas e expertise técnica.

Gestão de sistemas de cyber-infraestrutura institucional e arquitetura informacional. Decisões sobre arquitetura de informação e design são partes cruciais do projeto. Saber onde o projeto se situará institucionalmente, como ele será mantido, em quais ambientes de servidor e como o software e/ou as plataformas para o desenvolvimento de conteúdo serão escolhidas está na base do projeto. Também é necessário saber quem vai configurar a infraestrutura de servidor, administrar os sistemas, instalar o software (e mantê-lo atualizado) e fazer a cópia de segurança [*backup*] do conteúdo.

Softwares de código aberto e transferência de tecnologia. O desenvolvimento de ferramentas e plataformas é uma das pedras fundamentais dos projetos em Humanidades Digitais. É no interesse da empreitada comum de ensino e aprendizado que o software seja entendido como um recurso comunitário com seu código fonte compartilhado de forma a permitir o suporte e desenvolvimento pela comunidade de usuários como um todo. Em geral, projetos devem ser construídos com a visão de fomentar soluções comuns e plataformas compartilhadas, embora às vezes soluções pontuais servem propósitos específicos. Como o projeto permite a documentação e a transferência dos códigos, ferramentas, plataformas e aplicações?

Documentação. A documentação da estrutura e do design de um projeto é uma peça essencial do trabalho. Frequentemente isso é ignorado. A documentação é essencial para a continuidade do projeto após sua empreitada inicial, e é uma contribuição importante para o campo, bem como uma maneira de outros dar novos propósitos ao design. Processos de desenvolvimento devem ser documentados; especificações funcionais e técnicas devem ser documentadas; requisitos do sistema para o projeto devem ser documentados (por exemplo, quais browsers e versões são suportadas; quais plug-ins são requeridos); entidades de bases de dados e esquemas relacionais devem ser documentados; e, finalmente, o código deve ser documentado, incluindo as bibliotecas de código (*code libraries*) disponíveis publicamente utilizadas no projeto, acordos de licenciamento ou de usuários (especialmente para APIs – *Application Programming Interfaces*), e as operações pretendidas de módulos individuais, com atribuições de autor.

Considerações sobre audiência e usuários. Deixar claro quem é o público do projeto e como seus membros se engajam no seu desenvolvimento é importante, mesmo se a pesquisa é orientada pela curiosidade ou agenda de um acadêmico em particular. Projetos sem audiências ou usuários são silos nos quais o trabalho e os recursos desaparecem. O teste por usuários é geralmente uma parte crítica necessária ao refinamento da interface do projeto e das características de navegação.

Conformidade com todas as regulamentações legais. Projetos em Humanidades Digitais devem seguir os padrões do *Americans With Disabilities Act (ADA)* em seu design e devem estar em conformidade com as restrições de propriedade intelectual e direitos autorais. Este último, entretanto, deve ser aplicado com um claro entendimento do direito ao uso justo, o caráter sem fins lucrativos de quase toda a pesquisa em humanidades, e a contribuição que tal pesquisa faz para o conhecimento e o reconhecimento de objetos e heranças culturais.

Modelos de publicação e disseminação. Chamar a atenção para um projeto digital requer disponibilizá-lo para visualização em um sítio on-line, fazê-lo ser revisado e criar visibilidade dentro de uma comunidade acadêmica e entre os usuários potenciais e futuros colaboradores. Projetos devem ter um plano de disseminação e publicação. Projetos construídos com e a partir de comunidades despertam mais interesse do que projetos construídos por acadêmicos individualmente. Projetos digitais não devem “recriar a roda”, mas, ao invés disso, avaliar estrategicamente e, quando possível, tirar vantagens de soluções de software, plataformas e ferramentas existentes. Tanto o futuro do aprendizado em humanidades como a habilidade da produção acadêmica em humanidades de fazer a diferença na sociedade como um todo dependem da circulação irrestrita do conhecimento acadêmico; por conseguinte as licenças menos restritivas devem ser a norma.

Crítérios de avaliação. Um projeto deve ter suas métricas de sucesso e fracasso afirmadas explicitamente. Essas podem variar desde criar um projeto que prove um conceito ou demonstre um princípio de design a um projeto que estabeleça o objetivo de digitalizar e classificar um montante particular de material ou engajar uma comunidade específica no discurso e discussões on-line. Possuir objetivos e referenciais claros é útil como uma forma de avaliar a relação entre os recursos e os resultados.

Diálogos com múltiplos campos. O projeto dialoga com outros trabalhos no seu campo, tanto aqueles concebidos tradicionalmente como aqueles realizados em mídia digital? Os autores compreendem e fazem referências a outras pesquisas e projetos digitais como modelos? Como o projeto se situa dentro do desenvolvimento intelectual de um dado campo ou campos?

Sustentabilidade. Independentemente de quão experimental seja sua base tecnológica, estratégias de preservação são uma característica definidora do design de um bom projeto. Ativos digitais são frágeis por natureza e sua fragilidade precisa ser considerada desde o início por meio de uma estratégia de preservação em médio e longo prazos. Qual é o plano

para sustentar o projeto digital? Onde ele será alocado e mantido institucionalmente? Como serão tais recursos sustentados? Qual será o custo de continuar o projeto, se ele for permanente, e quais possíveis fontes de receita há para seu suporte? O trabalho das equipes, estudantes e consultores bem como os custos de hardware, software e outros materiais devem ser levados em conta, sem mencionar os compromissos intelectuais do pesquisador principal e da comunidade de conselheiros e colaboradores.

Transparência. Todas as fontes de financiamento, sejam elas monetárias ou doações em espécie, devem ser divulgadas nos vários resultados aos quais um projeto em Humanidades Digitais dá origem.

Especificação 3: competências essenciais em processos e métodos

Quais são as habilidades básicas essenciais para estar apto ao trabalho em Humanidades Digitais? Como podem tais projetos ser mantidos em um ambiente acadêmico ou institucional? Esta orientação lista os elementos fundamentais necessários à criação de projetos digitais de pesquisa. As competências específicas variam por campo e disciplina e não todos os projetos requerem todas essas competências.

Todos os projetos digitais têm aspectos técnicos, administrativos e intelectuais em sua produção. À medida que ferramentas e plataformas criadas especificamente para as Humanidades Digitais se tornam a cada dia mais disponíveis, construir projetos customizados somente se justificará se uma nova ferramenta ou plataforma for parte do desenvolvimento ou se o projeto possui comprovadamente elementos únicos que requeiram uma solução particular.

Técnicas. Desenvolvimento web, infraestrutura, ambientes de servidor, design de interface, escolhas de ferramentas, plataformas, software e hardware.

Familiaridade com os tipos de dados e formatos de arquivo. Sobre qual fundamento as decisões sobre formatos de arquivo e tipos de dados são feitas?

Conhecimento de bases de dados. Se uma base de dados é parte de uma arquitetura de informação, qual é o seu tipo? Como ela irá funcionar e porque ela é necessária? Quais são as entidades na base de dados, quais são seus atributos e relações, e como os objetos serão consultados e ordenados? A base de dados possui código aberto, proprietário e/ou licenciado? Quais conjuntos de dados serão usados no projeto e quem os controla? Que tipo de permissões e direitos governarão os conjuntos de dados?

Dados estruturados em XML. Que esquema ou versão de *XML* está sendo utilizado e por que? É ele usado para a marcação (*mark-up*) ou apenas para metadados?

Padrões de metadados. Qual processo de seleção de metadados foi utilizado e como o padrão de metadados se ajusta ao projeto e a seu campo disciplinar, bem como a seu lar institucional? Estão os padrões de metadados em conformidade com os padrões existentes no campo?

Linguagens de script. Em que extensão são utilizadas as linguagens de *script* no projeto e como elas são apropriadas ao servidor e ao ambiente administrativo no qual elas trabalham, bem como às tarefas para as quais são designadas?

Plataformas GIS e dados espaciais. Ferramentas para o mapeamento e a análise espacial têm sido desenvolvidas dentro de disciplinas geográficas para o uso profissional, mas outras ferramentas de mapeamento mais populares (como o Google Earth) têm um limite mais baixo para utilização. Quais são os aspectos espaciais (e temporais) dos dados e como esses dados serão apropriadamente classificados para a análise? Como eles serão exibidos em um sistema GIS (*Geographic Information System*) ou de mapeamento, e quais são as questões de pesquisa que podem ser testadas com tais sistemas? Os dados já são “espaciais” e, se não, é esse processo automatizado ou ele envolve a georretificação manual de materiais (sejam mapas, fotográficas históricas, vídeos ou histórias orais)? Como isso será feito, por quem e com vistas a quais padrões de visualização e compartilhamento dentro e através de aplicativos de geonavegação virtual (*geo-browser*)?

Ferramentas de simulação virtual. Mundos virtuais e modelagem tridimensional são ferramentas para criar ambientes imersivos para a pesquisa histórica e sua apresentação. Novamente, quais ferramentas, softwares e sistemas estão sendo utilizados e para quais fins? Quais padrões estão sendo seguidos e como várias comunidades de praticantes irão se engajar nos modelos, simulações e mundos virtuais? Em quais plataformas existentes serão os modelos alocados e que tipos de restrição essas plataformas têm?

Plataformas existentes e emergentes para gestão de conteúdo e autoria. Como o projeto administrará o conteúdo existente e suportará o crescimento de novo conteúdo? Quem são os autores desse conteúdo proposto e como eles vão inseri-lo? Eles vão precisar ser tecnicamente habilitados ou uma interface do browser vai permitir sua participação? Quais sistemas de gestão de conteúdo são utilizados na infraestrutura ou repositório? Os sistemas de gestão de conteúdo permitem que os dados sejam compartilhados através de plataformas e repositórios?

Design de interface como modelagem de conhecimento. Como o conteúdo é exibido na interface e como o usuário navega esse conteúdo por meio da interface? Qual é o modelo de interface e como ele expressa o modelo de conhecimento do projeto e dá suporte a sua missão?

Mecanismos de jogos. A economia de jogos tem um papel a desempenhar no trabalho acadêmico bem como no entretenimento. Compreender a maneira como mecanismos de jogos podem ser incorporados em um projeto para dar suporte à participação é útil em certas circunstâncias.

Design para a mobilidade e a diversidade. O projeto possui dimensões que tornarão seu conteúdo disponível em aplicativos móveis ou permite que ele seja redirecionado para uso em múltiplos contextos? O projeto funcionará em diferentes plataformas? Ele funcionará através de fronteiras culturais, linguísticas e sociais? Está o projeto em conformidade com o ADA (*Americans With Disabilities Act*) ou ele tem limitações de uso por pessoas com deficiências?

Customizado vs. pronto para uso. É parte da pesquisa do projeto o design e a construção de uma plataforma ou ferramenta, e, em caso afirmativo, pode esse trabalho ser reorientado ou generalizado a partir de seu uso customizado para um público mais amplo? Se soluções prontas para uso ou sistemas de software padrão estão sendo utilizadas, como eles foram escolhidas? Muitas vezes projetos em Humanidades Digitais serão uma combinação dessas duas abordagens, utilizando APIs existentes, sistemas padrão de gestão de conteúdo, ou mecanismos de *blogging* que podem ser customizados de várias formas e estendidos para atender necessidades específicas de um projeto.

Intelectuais. Enquanto o elemento intelectual mais visível é geralmente o conteúdo, é importante reconhecer que os projetos em Humanidades Digitais apresentam argumentos e experimentos de conhecimento de muitas maneiras diferentes, geralmente contribuindo para a criação de novo conhecimento por meio de complexas interações, visualizações, dados e estruturas de dados, e até mesmo códigos. Projetos em Humanidades Digitais não são somente sobre conteúdo (embora isso é geralmente primordial), mas também sobre o design de múltiplos níveis de conhecimento e argumento que vão das operações na base de dados interna (*back-end*) aos pontos de acesso externo (*front-end*) de uma interface de usuário.

Comunicação intercultural. As maneiras pelas quais o design do projeto trabalhará interculturalmente foram consideradas? Ele se propõe a engajar comunidades cuja língua e/ou orientação cultural são variadas?

Imaginação generativa. É o projeto generativo e vai ele continuar a criar novos conteúdos, diálogos, debates e engajamento, ou é ele majoritariamente um repositório empacotado de conteúdo destinado a ser visualizado e utilizado, mas não alterado por meio de contribuições ou extensões? Ambos os casos são valiosos e servem a diferentes necessidades, públicos e objetivos intelectuais.

Pensamento lateral e iterativo. Como pode o projeto mudar com o tempo, e como as reflexões sobre suas limitações serão utilizadas para melhorar cada iteração? Pode o projeto “interagir bem” com outros projetos compartilhando dados através de serviços web de estrutura ou de módulos de código através de repositórios de compartilhamento de códigos?

Administrativos. Alocação de recursos, orientações para relatórios, descrições claras de trabalho, objetivos e definições de responsabilidades para todos os envolvidos são cruciais e devem ser especificadas em, ao menos, um memorando de notificação.

Propriedade intelectual. Abrangência de direitos e de direitos autorais de propriedade intelectual foram geridos e documentados? Os termos para uso do conteúdo devem ser postados claramente no site, e a informação de contato para fazer perguntas sobre o uso da propriedade intelectual deve ser fácil de localizar.

Circunstâncias institucionais. Qual é o lar institucional para esse projeto e quem será responsável por sua manutenção depois de o projeto ser construído? Custos e impactos em recursos humanos e materiais devem ser avaliados.

Sustentabilidade, financiamento e preservação. Planos de longo prazo para sustentabilidade podem incluir a migração do projeto para um repositório institucional, ou o arquivamento em um servidor ou provedor de serviços pago, ou a criação de um fluxo de receitas e modelo de negócios para seu contínuo suporte e manutenção. A colaboração com entidades institucionais, particularmente bibliotecas e repositórios de dados, serão necessários para a preservação de dados criados para e por um projeto em Humanidades Digitais. Podem os dados ser “extraídos” facilmente do projeto e arquivados em formatos padrão que sejam amplamente legíveis? Que tipo de plano de gestão de dados foi criado e como ele será implementado? Há quaisquer preocupações concernentes à privacidade ou segurança que necessitem ser tratadas?

Especificação 4: resultados em aprendizado para as Humanidades Digitais

Enquanto os principais padrões de avaliação se mantêm contínuos com aqueles da pedagogia tradicional das humanidades, baseada na sala de aula, as Humanidades Digitais reconhecem a importância de resultados adicionais produzidos pelo aprendizado prático, experiencial, e baseado em projetos através do fazer. A pedagogia das Humanidades Digitais enfatiza o trabalho de equipe e implica um papel aumentado da avaliação por pares, bem como atenção a um conjunto mais amplo de habilidades além do pensamento crítico e comunicação baseada em textos. Resultados enfatizam a habilidade de pensar criticamente com métodos digitais para formular projetos que têm questões das humanidades em seu bojo. Entre os resultados de aprendizado nas Humanidades Digitais priorizamos os seguintes:

Habilidade de integrar objetivos, métodos e mídias de pesquisas conduzidas digitalmente com investigações de disciplinas específicas. Adquirir e demonstrar novas fluências a partir de trabalhar com e navegar através de várias plataformas de informação para conceitualizar e levar adiante pesquisa em disciplinas específicas. Na prática, isso significa juntar as ferramentas tradicionais do pensamento humanístico (interpretação e crítica, perspectiva histórica, análise comparativa cultural e social, contextualização, pesquisa de arquivos) com as ferramentas do pensamento computacional (design de informação, análise estatística, sistema de informação geográfica, criação de bases de dados e computação gráfica) para formular, interpretar e analisar o problema de pesquisa baseado em humanidades.

Habilidade de compreender, analisar e utilizar dados. Demonstrar habilidade de sintetizar dados de múltiplas fontes e aproveitar tecnologias multimodais e multimídia para produzir argumentos digitais. Criar capacidade de formular um problema ou questão de pesquisa que se preste à abordagem computacional. Desenvolver habilidade de analisar problemas aplicando métodos digitais a dados baseados nas humanidades e interpretar os resultados de análises digitais e resultados produzidos computacionalmente de uma maneira criticamente significativa.

Desenvolver habilidade crítica para avaliar fontes e dados. Julgar a confiabilidade da informação e do conhecimento apresentado em um ambiente digital requer habilidades de discernimento para examinar a fonte, a autoridade e a legitimidade do material digital. Em relação aos dados, isso significa examinar como eles foram obtidos, classificados, armazenados e tornados acessíveis de várias maneiras ao usuário final.

Habilidade de usar criticamente o design. Compreender a importância do design de conhecimento na comunicação, desenvolvimento do projeto e preservação a longo prazo de dados digitais de formas que vão além da competência de um entendimento crítico das ferramentas, seus usos e limitações. Desenvolver a habilidade de usar o pensamento em design computacional para produzir formas de argumento e expressões de interpretação.

Habilidade de avaliar criticamente a informação e as tecnologias de informação. Interrogar a informação digital, visual e multimodal como evidência e criticar sua formação e validade. Criticar as características digitais de publicações em relação a (a) relevância acadêmica, (b) melhores práticas (por exemplo, citação e referência online, transparência de recursos e dados), (c) atribuição, (d) autoridade e rigor argumentativo. Compreender e criticar as epistemologias, visões de mundo e estruturar premissas inseridas nas plataformas digitais, tecnologias, visualizações e até linguagens computacionais.

Habilidade de trabalhar colaborativamente. Pensar através de disciplinas, mídias e metodologias em projetos de pesquisa de multiautoria, propostas de projetos, relatórios, e apresentações destinadas tanto a comunidades acadêmicas como não acadêmicas. Trabalho em equipe e participação em avaliações por pares. Tomada de conhecimento do ciclo de vida de desenvolvimento de um projeto em Humanidades Digitais e a habilidade de compreender as necessidades e prioridades em cada fase do desenvolvimento. Atender prazos agressivos e produzir, de forma completa e totalmente funcional, protótipos digitais, produtos, ferramentas de pesquisa e publicações. Identificar e avaliar contribuições específicas e papéis em projetos colaborativos para propósitos de avaliação por pares e crédito intelectual.

Especificação 5: criando apoio

Entre outras atividades, a produção acadêmica digital se dirige à possibilidade de relações transformadas entre consumidores e produtores do trabalho cultural. A seguir estão listados um conjunto de considerações para refletir sobre a significância cultural do trabalho em humanidades de transformar indivíduos em prosumidores⁹ com um *insight* crítico sobre as plataformas digitais. Também contém um punhado de pontos cruciais sobre os quais argumentar em favor das Humanidades Digitais como um campo.

⁹ Nota da Tradutora. A palavra “prosumidor” é a versão em português do inglês “*prosumer*”, um neologismo formado pela junção das palavras “produtor” e “consumidor”, utilizada para referir-se ao papel mais ativo desempenhado pelos consumidores na sociedade contemporânea.

Valor da memória cultural. A produção acadêmica humanística está engajada com a produção, preservação e interpretação da memória cultural. Aferir o valor de materiais legados e verificar o valor de contribuições contemporâneas é essencial. De quais maneiras o projeto contribui para a memória cultural (através da preservação dos materiais, através de interações entre os colaboradores, por meio de modos de engajamento público e assim por diante)?

Valores humanísticos, significância cultural e legitimidade. Demonstrar o valor de métodos interpretativos e valores humanísticos fundamentais como um contraponto para aqueles da cultura gerenciada é uma parte essencial da argumentação. Como são os valores e perspectivas das humanidades uma parte central das contribuições do projeto? No que o projeto contribui para a memória cultural e como essa memória é legitimada (e por quem)?

Noções expandidas de comunidade e participação. Para quem esse projeto tem valor e como essas pessoas são engajadas em sua produção, recepção e preservação? Que noções de comunidade e participação são centrais para o projeto? Como a participação é disponibilizada, gerida e facilitada? Como são tomadas as decisões sobre permissão de participação, inclusão e/ou exclusão e quem as toma? Quais são os limites, as obrigações e desafios remanescentes para a participação sem restrições?

Habilidade de analisar modalidades de organização e apresentação. Habilidade de compreender as formas em que a mídia organiza e apresenta argumentos são os fundamentos do uso informado da informação em qualquer ambiente. As características específicas da mídia digital – em suas formas múltiplas, híbridas e superpostas – precisam de suas próprias linguagens de avaliação.

Consciência reflexiva sobre regimes coercitivos. Toda mídia esconde bem como revela as regras de acordo com as quais ela inclui certos tipos de expressão e proíbe outros. O que é possível em um dado espaço digital ou projeto, e o que não é? Devemos ser pensadores dialéticos, reflexivos, conscientes de que qualquer “solução” sempre proíbe certas questões e problemas de surgir, ao passo que privilegia aquelas mesmas para as quais ela é a resposta. Todas as tecnologias são coercitivas em algum aspecto, e muitas têm se tornado tão naturalizadas que não mais as consideramos coercitivas, mas, ao invés disso, auto evidentes e necessárias. É tarefa dos humanistas digitais desnaturalizar essas tecnologias e criar fissuras para possibilidades novas e imaginativas surgirem.

Pensar além das ideologias dos modelos e do discurso estruturado. Como interpretamos a corporificação da dinâmica do poder e das relações na organização de espaços e processos estruturados? O ambiente digital estrutura sua expressão ideológica nas interfaces gráficas, nos tipos de dados, nas relações de bases de dados, bem como no conteúdo de cada projeto. A desfamiliarização epistemológica – o “tornar estranho” é uma característica importante do pensamento crítico moderno. A força do deleite, surpresa e até alienação em face de invenções inovadoras são os elementos iluminadores do pensamento imaginativo contemporâneo. O que pode ser mostrado para nos despertar do nosso consumo passivo? E como novas formas de conhecimento, engajamento e design se tornam os próprios meios de provocar a investigação, gerar pensamento, aprofundar valores e contribuir para a memória cultural de nossa espécie?

De consumidor passivo a prosumidor ativo. O papel do leitor e visualizador varia daquele de um consumidor do material exibido para aquele do prosumidor criticamente informado e discriminador do material cultural. Como o projeto facilita o engajamento crítico e produtivo ao invés do consumo passivo?

Criação de cidadãos acadêmicos e acadêmicos cidadãos. Muitos projetos dão suporte à participação substantiva de amadores, acadêmicos sem afiliação profissional cuja expertise em um campo é altamente desenvolvida, informada e dirigida pela paixão intelectual. De quais formas o projeto integra (e também avalia) a multiplicidade de perspectivas e os criadores de conhecimento? Como os acadêmicos – tradicionalmente concebidos – se engajam com uma cidadania pública mais ampla, e, similarmente, como os cidadãos se engajam no projeto intelectual de criação de conhecimento como acadêmicos?

Acesso aberto para as máquinas

Christof Schöch¹

Tradução do alemão²: Winfried Nöth

Resumo: Muitos debates sobre o Acesso Livre giram atualmente em torno de modelos de financiamento adequados, perdendo de vista o fato de que os formatos de publicação também devem ser considerados como fundamentalmente mais abertos, ou seja, para além do formato PDF como a reencarnação digital do livro impresso, para que o potencial das tecnologias digitais de produção de conhecimento possa ser utilizado da melhor forma possível. Por conseguinte, o presente artigo trata da questão das publicações científicas como dados (mas não da publicação de dados de investigação). A exigência central do documento é trabalhar no sentido da substituição do formato PDF e do desenvolvimento e utilização de formatos de dados abertos e padronizados para publicações científicas que satisfaçam os princípios FAIR. O artigo argumenta que, para além dos metadados, das palavras-chave, da estrutura do texto e das referências bibliográficas, o conteúdo do texto é de particular importância: são necessários modelos e soluções técnicas para a forma como as declarações centrais de uma publicação científica podem ser incorporadas na própria publicação de uma forma legível por máquina. Para todos estes aspectos, formatos semiestruturados amplamente utilizados como XML (por exemplo TEI, JATS ou RDF) e JSON (como em BibJSON, semelhante também a BibTex), bem como o princípio da Web Semântica com a utilização de Dados Abertos Vinculados desempenham um papel importante.

Palavras-chave: Dados Abertos Vinculados. Acesso aberto. Publicação científica. XML. BibTex.

¹ Professor de Humanidades Digitais na Universidade de Trier, Alemanha, codiretor do Centro Trier de Humanidades Digitais, presidente da COST Action Distant Reading for European Literary History e da Associação de Humanidades Digitais nos países de língua alemã (DHD) dh.uni-trier.de.

² de Schöch (2020).

Open access to machines

Abstract: The debates about Open Access currently revolve around the question of suitable financing models, losing sight of the fact that publication formats must also be thought of as fundamentally more open. That means looking beyond the PDF format as the digital reincarnation of the printed book, if the full potential of digital technology to produce knowledge is to be exploited in the best possible way. This paper therefore deals with the question of scientific publications as data (but not with the publication of research data). The central demand of this paper is to work towards a replacement of the PDF format and the development and use of open, standardised data formats for scientific publications that meet the FAIR principles. The paper argues that, in addition to metadata, keywords, text structure and bibliographic references, text content is of particular importance. Models and technical solutions are needed for how central statements of a scientific publication can be embedded in the publication itself in a machine-readable form. For all these aspects an important role can be played by the use of semi-structured formats such as XML (for example TEI, JATS or RDF) and JSON (as in BibJSON, similar to BibTex), as well as the semantic web with the use of Linked Open Data.

Keywords: Linked Open Data. Open Access. Scientific publishing. XML. BibTex.

Debates atuais sobre acesso aberto

O debate sobre o Acesso Livre nas ciências e humanidades mudou consideravelmente nos últimos anos. Mesmo que a prática fique muitas vezes um pouco atrás das convicções e da medida em que os servidores de *preprints* são utilizados, a proporção de revistas de acesso aberto ou a disponibilidade gratuita dos anais de conferências varia, visto que em grandes partes do sistema científico, pelo menos para artigos de revistas e artigos de conferências, já não há qualquer dúvida de que a publicação em acesso aberto faz sentido e é cientificamente apropriada (ver open-access.net/en/open-access-in-individual-disciplines para as diversas disciplinas). Um dos argumentos mais comuns é que os resultados da investigação com financiamento público devem também estar disponíveis ao público em sentido lato; e que o progresso científico pode ser mais bem promovido pela livre disponibilidade de publicações científicas em todo o mundo.

O debate atual é, portanto, menos sobre o porquê do que sobre o como. Nisso, prevalece um enfoque claro sobre a questão dos modelos de financiamento adequados. A questão central é como é que os custos de publicação e divulgação, por um lado, e os custos de desenvolvimento e disponibilidade de infraestruturas de publicação a longo prazo, por outro, podem ser financiados se isso já não for feito através de subscrições, como é atualmente o caso. Entre os modelos atualmente discutidos e praticados estão os chamados Encargos de Processamento de Artigos (isto é, as taxas de publicação a serem cobradas dos autores, suas instituições ou patrocinadores do projeto), grandes acordos de leitura e publicação (tais como os contratos com as principais editoras em nível nacional que o consórcio DEAL está tentando concluir) ou novos modelos de financiamento coletivo (tais como o modelo de adesão à Biblioteca Aberta de Humanidades; <open-access.net/en/information-on-open-access/business-models>; SPEICHER *et al.* 2018). Como pode ser feita uma relocação dos orçamentos de assinatura por bibliotecas para a promoção de editoras e iniciativas que publicam em Acesso Livre? Como se pode evitar que a atual injustiça de acesso (apenas aqueles que podem pagar são autorizados a ler os resultados científicos de outros) seja meramente substituída por uma

injustiça de publicação (apenas aqueles que podem pagar são autorizados a publicar resultados científicos) (PIRON, 2020; POOLEY, 2020)? Esta última questão em particular tem também uma forte dimensão internacional e é, portanto, também de importância para a política de desenvolvimento.

A solução da questão do financiamento é sem dúvida de grande importância, não apenas, mas especialmente para as Humanidades. Entretanto, a presente contribuição terá que se concentrar num outro aspecto da questão – que atualmente está sendo negligenciado nos debates intensivos sobre o financiamento –, a saber, a questão dos formatos de publicação adequados para a ciência. Na prática, o arquivo PDF, que é o equivalente digital do livro impresso e do artigo de periódico, domina claramente. Três características explicam a aceitação e o sucesso deste formato: (1) correspondência direta entre a versão impressa e a versão digital até o layout; (2) preservação da paginação e, portanto, possibilidade de continuar as práticas de citação familiar; (3) a aparente imutabilidade e, portanto, confiabilidade de um arquivo PDF. No entanto, exceto para distribuição e leitura individual, este formato só é adequado até certo ponto (apesar de algumas extensões como PDF/A para arquivamento e PDF etiquetado para melhor acessibilidade). Neste contexto, o Grupo de Trabalho de Publicação Digital da Associação Alemã *Digital Humanities* recomenda, por exemplo: “não use o PDF como o formato primário de publicação (camada de codificação), mas sim, se for o caso, como um formato de leitura derivado” (DHD-AG, 2020). Tanto para o arquivamento a longo prazo como para a avaliação computadorizada de coleções maiores de publicações, outros formatos têm claras vantagens.

Publicações científicas como dados

Assim que se trata menos de um punhado de publicações da sempre crescente literatura de pesquisa, ou seja, se trata não mais apenas de uma questão de leitura, mas de usar as publicações como base de dados para uma análise quantitativa, os numerosos pontos fracos do formato PDF se tornam aparentes. Se os princípios FAIR (*Findable, Accessible, Interoperable, Re-useable*) forem aplicados às publicações científicas como dados em vez de à publicação de dados de pesquisa, torna-se rapidamente claro o quão desastroso é a prática atualmente dominante de publicar exclusivamente como arquivos PDF (WILKINSON; DUMONTIER; MONS, 2016). É verdade que tais contribuições podem ser encontradas (são *findable*) através de identificadores e metadados persistentes, que estão fre-

quentemente disponíveis atualmente. Publicadas em Acesso Livre, elas também são acessíveis (*accessible*) sem grandes obstáculos financeiros ou técnicos. Porém, elas também são só parcialmente interoperáveis e reutilizáveis (*interoperable* e *re-useable*). Assim, o texto em um arquivo PDF pode ser extraído, mas praticamente sem nenhuma informação estrutural essencial. A separação entre título corrido, texto principal e anotações é, na melhor das hipóteses, apenas acessível indiretamente, através de referências tipográficas ou outros padrões. Dentro do texto principal, não é possível distinguir de forma confiável entre diferentes seções de texto (por exemplo, resumo, introdução, parte analítica ou interpretativa, resultados, ou mesmo entre texto principal e citações em bloco). Também informações semânticas não estão explicitamente disponíveis, porque entidades (pessoas, obras, organizações) ou conceitos (termos técnicos ou abstratos) não podem ser abordados especificamente dentro do texto. Da mesma forma, não é possível pesquisar especificamente autores ou autoras, editores, títulos, datas de publicação ou organizadores dentro dos dados bibliográficos.

Nessas circunstâncias, as publicações científicas não podem fazer valer plenamente o seu potencial. Eles só podem fazê-lo se, além de serem publicados digitalmente e com livre acesso, também estiverem disponíveis em formatos estruturados e semanticamente enriquecidos. As estratégias de publicação correspondentes, que tanto quanto possível produzem não apenas publicações legíveis por humanos, mas também por máquinas, têm sido discutidas há dez anos (com base na ideia da Web Semântica) sob o título de *Semantic Publishing* (SHOTTON, 2009). Semelhante ao caso da construção e publicação de dados nas ciências humanas, grandes quantidades de publicações científicas são úteis, mas ainda melhor são publicações semanticamente e estruturalmente enriquecidas, que assim se tornam conjuntos de dados.

Alguns cenários de aplicação relevantes neste contexto estão resumidamente delineados a seguir. Por exemplo, a análise linguística da linguagem científica está interessada nas propriedades linguísticas de textos de diferentes disciplinas ou de diferentes tipos de literatura científica. Ela poderia lidar com dados de publicação anotados estruturalmente, mas também com vocabulário, estilística e padrões de argumentação de seções funcionalmente diferentes de textos científicos (tais como introdução, parte principal ou conclusão). A pesquisa quantitativa sobre a história de uma disciplina, por exemplo, poderia ganhar uma base empírica muito mais ampla através de análises detalhadas e em larga escala de redes

de citação baseadas em bibliografias estruturadas. E a pesquisa baseada em um *corpus* de textos sobre um determinado autor, obra ou problema, que já se beneficia de resumos e palavras-chave, poderia funcionar muito mais precisa e extensivamente, se as entidades essenciais no texto fossem anotadas e as declarações centrais de todas as publicações fossem legíveis por máquina e automaticamente vinculáveis a uma rede de declarações.

Os requisitos essenciais para as possibilidades que tais publicações científicas legíveis por máquina devem oferecer podem ser resumidos da seguinte forma:

1. Codificação estruturada e padronizada de metadados relacionados a documentos (incluindo informações bibliográficas; palavras-chave; licenças; identificadores persistentes, tais como DOIs);
2. Codificação explícita das estruturas de texto (entre outros, texto principal versus anotações; introdução, parte principal, conclusão; se necessário, dados, hipóteses, métodos, resultados; texto do autor versus citações);
3. Codificação estruturada de referências bibliográficas (incluindo informações bibliográficas, incluindo identificadores persistentes, como DOIs para literatura de pesquisa e, quando apropriado, fontes primárias);
4. Rotulagem legível por máquina das entidades (atores, organizações, lugares, horários) e conceitos em uma contribuição (resumo, termos técnicos);
5. Representação legível por máquina das declarações centrais de uma contribuição.

Os benefícios das quatro primeiras exigências aqui mencionadas são, em grande parte, indiscutíveis. Predominantemente, existem também soluções técnicas que simplesmente precisam ser utilizadas e, para promover este uso, (melhor) apoiadas pelas infraestruturas de publicação existentes ou mais abertas para uso posterior por terceiros.

A codificação estruturada dos metadados relacionados a documentos (requisito 1) é atualmente tratada, de maneira predominante, separadamente dos próprios textos dos artigos nas bases de dados dos fornecedores, onde são, naturalmente, utilizados intensivamente para fins de descoberta. Uma melhor integração poderia ser implementada através da

incorporação dos metadados correspondentes na área “Propriedades” de um arquivo PDF. Outros métodos são manter o DOI como referência para a contribuição e os metadados correspondentes ou, se for utilizado um formato semiestruturado, codificar esses metadados na área correspondente do arquivo XML (por exemplo, em TEI ou JATS <jats.nlm.nih.gov/; e em TEI: tei-c.org>).

No caso da codificação explícita de estruturas de texto, por exemplo, atribuindo seções de texto a classes estruturais ou semânticas (requisito 2), ocorre que isto geralmente não é implementado no contexto de arquivos PDF (apesar das possibilidades realmente disponíveis). No formato PDF domina demais o aspecto do layout. As infraestruturas que poderiam utilizar tais informações não estão suficientemente desenvolvidas. Aqui, somos dependentes ainda das possibilidades de formatos baseados em XML como JATS ou TEI, que até agora desempenham só um papel marginal no setor de periódicos. Até o momento, apenas pouquíssimas revistas ou editoras aceitam manuscritos em tais formatos. As exceções a esta regra são o *Digital Humanities Quarterly* (DHQ) e o *Journal of the Text Encoding Initiative* (jTEI), que usam XML-TEI (HARRISON, 2016). O fornecedor de periódicos *Public Library of Science* (PLOS) assim como *Open Library of Humanities* geram uma versão XML em JATS para seus artigos de periódicos e a oferecem para download. O LaTeX é frequentemente aceito nas áreas das Ciências Naturais e da Informática. Elsevier, por exemplo, o converte internamente para XML, mas não o publica nesse formato.

Para a codificação estruturada das referências bibliográficas (requisito 3), por outro lado, existe toda uma gama de formatos de dados bem estabelecidos, entre os quais a BibTex certamente provou ser particularmente central. Numerosas ferramentas, tais como Citavi (acesso pago; <citavi.com> ou Zotero (gratuito; <zotero.org>) permitem o gerenciamento conveniente de tais dados, bem como sua utilização na redação de textos científicos. Hoje, porém, estes formatos e programas são usados principalmente para gerar uma bibliografia uniformemente formatada num estilo de citação específico. A bibliografia é então anexada ao texto, mas com perda da estruturação explícita dos dados. Além disso, conceitos existentes para a extensão de tais dados, por exemplo, por informações sobre o propósito de uma referência em uma publicação usando uma ontologia como a Citação Tipo Ontologia (CiTO; <zotero.org>), são pouco utilizados. Para fazer melhor uso desses dados, são necessárias adaptações infraestruturais de longo alcance, como a possibilidade de adicionar um arquivo BibTex com os dados bibliográficos a uma determinada publicação como suplemento. Também revistas, que aceitam a submissão de artigos em LaTeX + BibTex geralmente publicam em PDF.

Dados Abertos Vinculados para codificação de conteúdo

Passemos agora aos dois últimos requisitos, que estão diretamente relacionados com o uso de Dados Abertos Vinculados. Desde o artigo de Shotton, a indústria editorial mudou drasticamente. No entanto, o que ele disse na época ainda se aplica: “com algumas exceções brilhantes, as revistas on-line atualmente não fornecem nenhuma margem semântica de texto que facilitaria uma maior compreensão do significado subjacente” (SHOTTON, 2009, p. 87). Os exemplos mencionados por Shotton não foram aceitos. A virada semântica da publicação científica ainda está por vir. Há certamente muitas razões para isto, entre elas provavelmente uma falta de consciência dos benefícios e das possibilidades de implementar tal codificação semântica. Este é o ponto do qual a presente contribuição gostaria de partir.

A leitura automática da indexação das entidades e conceitos de uma contribuição (requisito 4) não é algo inteiramente novo. Afinal, ela já é a base para a criação de um índice ou registro de palavras-chave, como é usual para livros de não ficção, geralmente com respeito a entidades (tais como pessoas, organizações, lugares e títulos de trabalho) por um lado, e a conceitos (resumos, conceitos, termos técnicos) por outro. O que é novo no contexto das publicações digitais, entretanto, é que a indexação não apenas conecta o registro dentro de uma publicação com as referências no texto e assim torna a publicação acessível, mas que as entidades e conceitos podem ser identificados de forma única, ligando-os a dados de controle de autoria (pt.wikipedia.org/wiki/Controle_de_autoridade) e integrados numa ontologia específica do domínio, e assim tornando-se parte da Web Semântica como Dados Vinculados (Abertos) (Linked [Open] Data) (DENGEL, 2012). Além disso, tais marcações devem ser inseridas não apenas manualmente pelos autores, mas devem também ser usadas ferramentas disponíveis para anotação automática (por exemplo, através do reconhecimento da entidade nomeada) e para integração em ontologias relevantes.

Tal integração do conteúdo do artigo na Web Semântica não só torna possível a indexação através de numerosas publicações, mas as entidades e conceitos indexados também podem ser dinamicamente enriquecidos com informações adicionais: pessoas mencionadas, por exemplo, por dados de vida e local de ação ou disciplina(s). Tudo isso requer infraestruturas em um sentido múltiplo: no sentido de formatos de dados que permitam um enriquecimento correspondente das publicações; de conjuntos

de dados padrão que possam ser referenciados para a desambiguação e enriquecimento das entidades; e de infraestruturas de publicação que permitam a indexação, vinculação e utilização dos dados correspondentes. No que diz respeito aos formatos de dados, o JATS é limitado em sua expressividade, enquanto o TEI fornece todos os mecanismos essenciais. No que diz respeito à infraestrutura de publicação, o autor muitas vezes não tem conhecimento de nenhuma plataforma de publicação, editoras ou periódicos que aceitem os formatos de dados anotados correspondentes no momento da submissão e que, assim, também utilizariam os dados para publicação. No entanto, há muito a aprender de outras áreas de aplicação de dados normativos, como por exemplo, na edição filológica (STADLER, 2012; KAMZELAK, 2016). Por último, mas não menos importante, a integração na Web Semântica também requer acesso aberto para que o acesso livre às publicações relevantes possa ser feito em todos os locais de publicação e não limitado apenas ao portfólio de um fornecedor. Entretanto, isto está em conflito direto com os interesses dos editores que querem manter seus leitores em sua própria plataforma.

O último requisito mencionado acima é que uma publicação legível por máquina deve oferecer suas principais declarações ou resultados de forma cuidadosamente modelada semanticamente. Seringhaus e Gershtein (2007) chamam isto de “Resumo Digital Estruturado” (*Structured Digital Abstract*), o que definem como um “resumo XML legível por máquina de fatos pertinentes no artigo”. Ao contrário das exigências discutidas até agora, esta exigência é geralmente menos aceita, pelo menos no contexto da redação de publicações científicas. Este estado de desenvolvimento está relacionado ao fato de que existem soluções técnicas menos específicas e suficientemente avançadas e que o tema em si tem sido conceitualmente muito menos bem refletido. Deve ser feita uma distinção entre a implementação técnica, por um lado, e a solução conceitual, por outro. A implementação técnica parece secundária em relação ao estado atual do debate. Ela é principalmente uma questão de construção de consenso e das ferramentas disponíveis em uma comunidade. Parece claro, entretanto, que tal implementação (como no caso de entidades e conceitos de rotulagem) deveria utilizar os mecanismos de *Linked Open Data* (LOD) e, portanto, da Web Semântica.

O enfoque, no que segue, será, portanto, no lado conceitual, provavelmente o aspecto mais controverso do tópico. No contexto literário, parte da dificuldade decorre também do seguinte: diferente da situação na biologia ou química, onde já foram desenvolvidas várias ontologias relevantes, ou da linguística, onde já existe uma vasta experiência e projetos

relevantes com a “Linguistic Linked Data Cloud” (linguistic-lod.org), na área do estudo de literatura, a utilização do Linked Open não está ainda muito enraizada – pelo menos não para além da codificação de metadados bibliográficos básicos e no desenvolvimento de edições digitais. Por esta razão, seguirão algumas reflexões nesta direção, utilizando um estudo de caso do campo da especialização do autor, a história literária.

O estudo de caso toma a perspectiva do enriquecimento retrospectivo de publicações científicas existentes por meio de *Structured Digital Abstracts*. A experiência a ser reunida aqui também será útil para responder à questão de como documentar o conteúdo de publicações científicas emergentes de uma forma legível por máquina. Em termos de conteúdo, o foco é o romance francês da segunda metade do século XVIII. Com base em uma bibliografia de todos os romances publicados na França entre 1750 e 1799 (existem cerca de 2.000 títulos diferentes) já modelada como um LOD (LÜSCHOW, 2020), as entidades (romances e romancistas) aí contidas foram enriquecidas com afirmações relevantes à história da literatura, assim sendo já acessíveis. As seguintes considerações fazem parte das considerações preliminares para o projeto *Mining and Modeling Text* (MiMoText), que está atualmente começando no Centro Trier de Humanidades Digitais da Universidade de Trier, coordenado pelo autor deste artigo e financiado no âmbito da iniciativa de pesquisa do estado da Renânia-Palatinado (kompetenzzentrum.uni-trier.de/de/projekte/projekte/m).

Um breve trecho de texto sobre o romance *Candide*, a partir de uma visão geral da história da literatura francesa de Erich Köhler, servirá de exemplo:

Candide é a obra mais lida de Voltaire e provavelmente já o foi durante a vida do autor. Quando foi impresso pela primeira vez em Genebra em 1759, foi imediatamente proibido, mas somente com o resultado de que houve treze novas edições no mesmo ano. (KÖHLER, 1984, p. 8)

A ideia básica agora é registrar o conteúdo central de um tal texto na forma de proposições basais, formuladas em forma de declarações “sujeito-predicado-objeto” no sentido de Linked Open Data, registradas como os chamados triplos, por exemplo, em um formato como RDF ou Turtle (ver DENGEL, 2012).

Antes de tudo, uma codificação semântica e explícita das declarações em uma publicação científica pode ser baseada nas entidades e conceitos anotados cuja distinção no sentido do requisito 4 (ver acima) é pressuposta aqui. Portanto, com a abordagem de Linked Open Data, já existem mecanismos para se referir a entidades e conceitos e utilizá-los

como entidades em declarações. No estudo de caso aqui tratado, as entidades relevantes são então pessoas (concretamente: autores de romances ou de literatura técnica; aqui: Voltaire) bem como obras (especificamente: romances, ou artigos individuais, capítulos, ou publicações monográficas; aqui: *Candide*). O inventário das entidades concebíveis deve ser entendido como uma lista incompleta e, portanto, não deve ser codificado. Além disso, entidades conceituais também podem ser derivadas de campos literários básicos, tais como conteúdo, estilo, gênero, época etc. A anotação (em XML-TEI e usando identificadores de conjuntos de dados padrão como o VIAF ou o *Getty Thesaurus of Geographical Names*) poderia então ter a seguinte forma:

```
<p><title type="work" ref="viaf:176620251">Candide</title> é a obra mais lida <persName type="author" ref="viaf:36925746"> de Voltaire</persName> e provavelmente já o foi durante a vida do autor. Quando <date>1759</date> foi lançado em <placeName type="city" ref="tgn:7007279">Genebra </placeName>, foi imediatamente banido, mas apenas com o resultado de ter sido reimpresso treze vezes no mesmo ano.</p>
```

No entanto, existem trabalhos exploratórios menos fundamentados conceitualmente, que podem ser utilizados para formular entidades como essas. Naturalmente, esses trabalhos podem, em primeiro lugar (e trivialmente), simplesmente conectar o próprio texto com as entidades mencionadas no texto, tal como

```
Köhler_1984 (viaf:174648806) HAS_SUBJECT Voltaire (viaf:36925746); Candide (viaf:176620251); Genebra (tgn:7007279)
```

Porém, com base nisso, a questão central é agora como o conteúdo do texto pode ser formalizado. Esta pergunta toca na compreensão fundamental de uma determinada disciplina, já que se trata de determinar que tipo de declaração uma comunidade científica de especialistas considera fundamental para um determinado domínio. O exemplo acima contém uma série de afirmações que seriam questionadas aqui e que são chamadas de afirmações pseudoformalizadas:

- Voltaire (viaf:36925746) IS_CREATOR_OF Candide (viaf:176620251)
- Candide (viaf:176620251) HAS_PUBLICATION_DATE 1759

- Candide (viaf:i76620251) HAS_PUBLICATION_LOCATION Genebra (tgn:7007279)
- Candide (viaf:i76620251) HAS_RECEPTION_INTENSITY alta
- Candide (viaf:i76620251) HAS_RECEPTION_TIME imediato;a longo prazo
- Candide (viaf:i76620251) HAS_LEGAL_STATUS censurado (1759)

Neste exemplo, os três primeiros itens não são muito mais do que metadados bibliográficos, tais como as que se encontram em catálogos ou bibliografias temáticas (e como já existem no nosso caso). Para alguns destes tipos de declarações, especialmente quando se trata de informação prosopográfica e bibliográfica, pode-se utilizar para a formalização ontologias existentes, por exemplo, Dublin Core (para creator, publisher, date, title, subject) ou as ontologias SPAR (para posterior modelação bibliográfica) (PERONI; SHOTTON, 2018; ver en.wikipedia.org/w/index.php?title=Dublin_Core&oldid=922336659). No entanto, isto não vale para as declarações seguintes. A questão central é, portanto, como deve ser concebida uma ontologia de tipos de afirmação central para um determinado domínio científico (aqui: história literária como parte dos estudos literários) e como pode ser estabelecido um consenso sobre estes temas na comunidade dos pesquisadores. Quais informações de domínio (aqui: informações da história literária) deviam ser formuladas como declarações básicas?

Comparativamente indiscutível, semelhante às declarações bibliográficas básicas já mencionadas, deviam ser informação estabelecidas prosopográficas como a que se encontra na Wikidata, por exemplo:

- (pessoa) DATE_OF_BIRTH (data); HAS_DATE_OF_DEATH (data)
- (pessoa) OCCUPATION (profissão)
- (pessoa) RELIGION (religião)
- (pessoa) MOVEMENT (ideologia, visão de mundo, crença)

Declarações um pouco mais específicas do domínio, ainda oficialmente padronizadas dentro de uma ontologia, mas praticadas, por exemplo, na Wikidata, são as seguintes (wikidata.org/wiki/Wikidata:List_of_properties/work#Wikidata_property_related_to_works_of_fiction):

- (pessoa) INFLUENCED_BY (pessoa)
- (pessoa) AWARD_RECEIVED (prémio)
- (obra) GENRE (gênero)
- (obra) CARACTERES (nomes de personagens)
- (obra) NARRATIVE_LOCATION (Localização geográfica)
- (obra) SET_IN_PERIOD (período de tempo)
- (obra) DERIVATIVE_WORK (obra)
- (obra) INSPIRED_BY (obra)
- (obra) NARRATOR (nomes de personagens)

Aqui começa a ficar evidente que um exame sistemático deste tipo de declarações sob a forma de uma ontologia ainda está pendente. Por exemplo, “RELIGION” e “MOVEMENT” têm uma relação pouco clara com “INFLUENCED_BY” no nível da pessoa e com “INSPIRED_BY” no nível do trabalho. Para alguns aspectos, as taxonomias ou ontologias existentes poderiam ser reutilizadas, por exemplo, na área de títulos profissionais (históricos e atuais), por exemplo, conforme o sistema HISCO (*Historical International Standard Classification of Occupations*; ver VAN LEEUWEN; MAAS; MILES, 2004). Para outros aspectos, tais como gêneros literários, épocas, formas ou temas, não há recursos comparativamente formalizados e consensuais. Evidentemente, os predicados usados no Wikidata até hoje também não são de maneira alguma suficientes para uma descrição histórica adequada de obras literárias, autores e épocas. Somente com relação a obras literárias, por exemplo, as seguintes informações adicionais seriam relevantes:

- (obra) HAS_FORM (prosa|verso|outro)
- (obra) HAS_NARRATIVE_PERSPECTIVE (autodiegético|homodiegético|heterodiegético) – forma narrativa em obras narrativas
- (obra) HAS_DIALOGUE_PROPORTION (porcentagem) – Porcentagem de discurso direto em um trabalho narrativo, como uma porcentagem de palavras ou frases.
- (obra) HAS_STAGE_DIRECTIONS (porcentagem) – Proporção de direções de cena num trabalho dramático, em porcentagem de palavras

Evidentemente, esta lista está longe de ser conclusiva. Uma modelagem sistemática do domínio ainda está pendente. Entretanto, deve-se acrescentar, neste contexto, que as informações coletadas ou extraídas não são consideradas fatos, mas sim afirmações. Na medida em que cada afirmação é atribuída a uma fonte, ela representa a opinião de um especialista ou do estado da pesquisa no momento da publicação. Consequentemente, um sistema de informação que coleta um grande número de tais declarações também pode conter declarações contraditórias ou incompatíveis, sem que o inventário do sistema seja considerado inconsistente.

O projeto atual “Mining and Modeling Text” (MiMoText) trata apenas indiretamente da questão de como as publicações científicas devem ser acessadas no futuro. Entretanto, o objetivo central é identificar e modelar semanticamente um certo inventário de tipos de declaração histórica literária em um *corpus* de literatura especializada existente (especialmente mais antiga, mais concisa) seguindo a abordagem delineada e, em grande parte, automaticamente. A historiografia literária mais antiga também deve ser tornada visível novamente para ser utilizável em larga escala através da modelagem e publicação como Linked Open Data, usando ontologias específicas do domínio. Além disso, informação sobre romances raramente lidos deve ser pesquisada desta forma e incorporada ao sistema de informação da história literária resultante. Desde que uma quantidade suficientemente grande de literatura especializada indexada semanticamente esteja disponível, isto criará um sistema de informação bibliográfica-histórica que suporta uma gama de cenários de aplicação. Por exemplo, será possível determinar o histórico de recepção de uma determinada autora não só quantitativamente (por exemplo, através do número de publicações relevantes por ano), mas também para compreender o conteúdo da obra, analisando a evolução dos respectivos temas abordados, as tendências de avaliação ou os respectivos autores comparativos mobilizados. Também será possível, com base nas declarações contidas no sistema (em termos de conteúdo, estilo, avaliação, classificação etc.), identificar obras literárias que, de acordo com os critérios escolhidos em cada caso, tenham características comuns e, portanto, possam ser adequadas para uma análise posterior e comparativa.

Esta forma de indexação semântica bastante elaborada e retrospectiva tornar-se-ia supérflua no futuro se a literatura especializada recentemente publicada já publicasse suas declarações essenciais sob a forma de Linked Open Data. A identificação das entidades relevantes e a formulação das declarações correspondentes poderiam ser feitas pelos próprios autores, ou poderiam ser desenvolvidos procedimentos que o fizessem automaticamente com base no texto completo, o que poderia ser feito

utilizando métodos desenvolvidos há algum tempo no campo da mineração de argumentos e da anotação semântica automática. No futuro, uma maior precisão (com simultaneamente menos cobertura) poderia ser esperada se uma prática correspondente de escrever publicações científicas fosse estabelecida, semelhante à prática atual de especificar uma série de palavras-chave ou selecionar termos de uma ontologia científica ao submeter artigos ou capítulos. Entretanto, ainda há um longo caminho a percorrer até lá, e os futuros métodos de publicação científica e a indexação retrospectiva da história das áreas de pesquisa terão um papel nisso. Por um lado, uma prática contemporânea de anotação semântica de publicações científicas em combinação com o texto completo também produz dados de treinamento para o desenvolvimento de procedimentos automáticos. Por outro lado, a indexação retrospectiva também pode ajudar a especificar os requisitos para futuros procedimentos de anotação semântica de publicações científicas e para ontologias subjacentes, específicas do domínio.

De qualquer forma, a visão do autor é que, num futuro próximo, não formularemos mais resultados de pesquisa apenas em prosa, em linguagem natural. Não produziremos, distribuiremos e receberemos mais artigos ou livros como arquivos PDF. Esta prosa também não será sem conexão com a publicação associada de conjuntos de dados e o código de programação. Em vez disso, este texto em prosa será vinculado a códigos e conjuntos de dados relevantes, fornecido com metadados ricos, marcado em sua estrutura de texto, anotado com entidades e conceitos usando dados bibliográficos estruturados, e publicado na forma de declarações LOD. Não se pretende alegar que o texto contínuo formulado em prosa em linguagem natural se tornará obsoleto como resultado. Porém, no futuro, o texto contínuo não ficará mais sozinho – ele será incorporado em um contexto rico, legível por máquinas de dados, código, com metadados, dados de citação e declarações modeladas. (Um suplemento de dados para esta publicação foi arquivado em [Zenodo.org](https://zenodo.org/10.5281/zenodo.3898418) sob o seguinte DOI: doi.org/10.5281/zenodo.3898418).

Referências

DENGEL, Andreas (org.). *Semantische Technologien: Grundlagen, Konzepte, Anwendungen*. Heidelberg: Spektrum, 2012.

DIGITAL HUMANITIES im deutschsprachigen Raum (org.). Digitales Publizieren. Digitales Publizieren / DHd-Arbeitsgruppe “Digitales Publizieren”. Wolfenbüttel: Herzog August Bibliothek, 2016. Disponível em: dhd-wp.hab.de/?q=content/working-paper-digitales-publizieren. Acesso em: 4 jun. 2020.

HARRISON, Melissa. Collecting XML at article submission at eLife: two steps forward, one step back? *In: Journal Article Tag Suite Conference (JATS-Con). Proceedings 2016* [Internet]. Bethesda (MD): National Center for Biotechnology Information (US), 2016. Disponível em: <ncbi.nlm.nih.gov/books/NBK350147>. Acesso em: 5 jun. 2020.

KAMZELAK, Roland S. Digitale Editionen im Semantic Web: Chancen und Grenzen von Normdaten, FRBR und RDF. *In: RICHTS, Kristina; STADLER, Peter (org.). „Ei, dem alten Herrn zoll' Ich Achtung gern“: Festschrift für Joachim Veit zum 60. Geburtstag.* München: Allitera, 2016, p. 423-435. Disponível em: <dx.doi.org/10.25366/2018.29>. Acesso em: 5 jun. 2020.

KÖHLER, Erich. *Vorlesungen zur Geschichte der Französischen Literatur: Aufklärung II*, org. RIEGER, Dietmar. Stuttgart: Kohlhammer, 1984.

LIPPI, Marco; TORRONI, Paolo. Argumentation mining: State of the art and emerging trends. *ACM Transactions on Internet Technology* v. 16, n. 2, p. 1-25, 2016. Disponível em: <doi.org/10.1145/2850417>. Acesso em: 3 jun. 2020.

LÜSCHOW, Andreas. *Bibliographie du genre romanesque français 1751-1800 – RDF Model.* Zenodo 2019. Disponível em: <doi.org/10.5281/zenodo.3401428>. Acesso em: 3 jun. 2020.

_____. Automatische Extraktion und semantische Modellierung der Einträge einer Bibliographie französischsprachiger Romane. *Jahrestagung des DHd-Verbands 2020: Spielräume.* Paderborn, 2020. Disponível em: <doi.org/10.5281/zenodo.3666689>. Acesso em: 3 jun. 2020.

PERONI, Silvio; SHOTTON, David. The SPAR ontologies. *In: VRANDECIC, D. et al. (org.) The Semantic Web – ISWC 2018. ISWC 2018. Lecture Notes in Computer Science*, vol. 11137, 2018. Cham: Springer. Disponível em: <doi.org/10.1007/978>. Acesso em: 3 jun. 2020.

PIRON, Florence. Qui sait ? Le libre accès en Afrique et en Haïti. 2020. Disponível em: <laviedesidees.fr/Qui-sait.html>. Acesso em: 3 jun. 2020.

POOLEY, Jeff. The library solution: How academic libraries could end the APC scourge, 2020. Disponível em: <items.ssrc.org/parameters/the-library-solution-how-academic-libraries-could-end-the-apc-scourge/>. Acesso em: 4 maio 2020.

SCHÖCH, Christof. Open Access für die Maschinen. *In: EFFINGER, Maria; KOHLE, Hubertus (org.). Die Zukunft des kunsthistorischen Publizierens.* *In: arthistoricum.net*, 2020. DOI: <doi.org/10.11588/arthistoricum.663.c9210>. Acesso em: 4 maio 2020.

SERINGHAUS, Michael R.; GERSTEIN, Mark B. Publishing perishing? Towards tomorrow's information architecture. *BMC Bioinformatics* v. 8, n. 17, 2007. Disponível em: <doi.org/doi:10.1186/1471-2105-8-17>. Acesso em: 3 jun. 2020.

SHOTTON, David. Semantic publishing: The coming revolution in scientific journal publishing. *Learned Publishing* v. 22, p. 85-94, 2009. Disponível em: <doi.org/10.1087/2009202>. Acesso em: 3 jun. 2020.

SPEICHER, Lara; ARMANDO, LORENZO; BARGHEER, Margo; EVE, Martin Paul; FUND, Sven; LEÃO, Delfim; MOSTERD, Max; PINTER, Frances; SOUYIOULTZOGLOU, Irakleitos. *OPERAS Open Access Business Models White Paper*, 2018. Disponível em: <doi.org/10.5281/zenodo.1323707>. Acesso em: 3 jun. 2020.

STADLER, Peter. Normdateien in der Edition. *Editio: Internationales Jahrbuch für Editionswissenschaft*, v. 26, n. 1, 2012. Disponível em: <doi.org/10.1515/editio-2012-0013>. Acesso em: 5 jun. 2020.

UREN, Victoria; CIMIANO, Philipp; IRIA, José; HANDSCHUH, Siegfried; VARGAS-VERA, Maria; MOTTA, Enrico; CIRAVEGNA, Fabio. Semantic annotation for knowledge management: Requirements and a survey of the state of the art. *Web Semantics: Science, Services and Agents on the World Wide Web* v. 4, p. 14-28, 2005.

VAN LEEUWEN, Marco H. D., MAAS, Ineke, MILES, Andrew. Creating a Historical International Standard Classification of Occupations: An exercise in multinational interdisciplinary cooperation. *Historical Methods: A Journal of Quantitative and Interdisciplinary History* v. 37, n. 4, p. 186-97, 2004. Disponível em: <doi.org/10.3200/HMTS.37.4.186-197>. Acesso em: 5 jun. 2020.

WILKINSON, Mark D.; DUMONTIER, Michel; MONS, Barend. The FAIR Guiding principles for scientific data management and stewardship. *Scientific Data* v. 3, n. 160018, 2016. Disponível em: <nature.com/articles/sdata201618>. Acesso em: 5 jun. 2020.

Researching visual semiotics online

Jamin Pelkey¹

Abstract: Analyzing visual meaning online and curating digitized images are topics of increasing relevance, but many potential methodologies for doing so remain merely implicit, underthematized, or unexplored. The potential for testing and developing semiotic theory through the exploration of visual data online also requires far more careful attention. In response, this paper provides an integrated, reflexive, Peircean account of two case studies featuring research projects focused on visual data drawn primarily from sources online, relying heavily on Google Image Search as a data collection tool. The first study illustrates the comparative analysis of brand mark logos to test and refine a theory of embodied semiotics involving oppositional relations. The second study illustrates the comparative analysis of images depicting the Tibetan Wheel of Life and Yama the monster of death, in order to test the embodied grounding hypothesis for the semiotic square. Issues of hypothesis formation, research parameters, data collection, database construction, operationalization, coding parameters, open data archiving and related issues are addressed in order to further develop and encourage practices of researching visual semiotics online in the context of Digital Humanities scholarship.

Keywords: Mixed-methods research. Google Image search. Visual content analysis. Semiotic theory. Semiotic methods. Peircean semiotics.

¹ Ryerson University. E-mail: jpelkey@ryerson.ca.

Pesquisando semiótica visual online

Resumo: Analisar significado visual online e realizar a curadoria de imagens digitalizadas são tópicos de crescente relevância, mas muitas metodologias potenciais para fazê-lo permanecem meramente implícitas, subtematizadas ou inexploradas. O potencial para testar e desenvolver a teoria semiótica por meio da exploração de dados visuais online também requer atenção muito mais cuidadosa. Em resposta, este artigo fornece um relato integrado, reflexivo e peirceano de dois estudos de caso que apresentam projetos de pesquisa focados em dados visuais extraídos principalmente de fontes online, baseando-se grandemente na Pesquisa de Imagens do Google (*Google Image Search*) como ferramenta de coleta de dados. O primeiro estudo ilustra a análise comparativa de logotipos de marcas para testar e refinar uma teoria da semiótica corporificada envolvendo relações de oposição. O segundo estudo ilustra a análise comparativa de imagens que retratam a Roda da Vida Tibetana e Yama, o monstro da morte, a fim de testar a hipótese fundamental da corporificação para o quadrado semiótico. Questões de formação de hipóteses, parâmetros de pesquisa, coleta de dados, construção de banco de dados, operacionalização, parâmetros de codificação, arquivamento de dados abertos e questões relacionadas são abordadas a fim de desenvolver e encorajar práticas de pesquisa semiótica visual online no contexto dos estudos acadêmicos em Humanidades Digitais.

Palavras-chave: Pesquisa de métodos mistos. Pesquisa de imagens do Google. Análise de conteúdo visual. Teoria semiótica. Métodos semióticos. Semiótica peirceana.

Introduction

The open, comparative nature of semiotic inquiry is well suited to massively interconnected, collectively constructed online world. In this virtual realm of overlapping semiospheres, context, connotation, and content are constantly shifting and visual layers of meaning negotiation prevails. Digital images, text, emoji, video, animation, layout, hypertext, and interface all interact and integrate relative to the habituated experiential aims of embodied interpreters, who are themselves acting and reacting in dialogue with the designs and ends of content creators and the interests of their financial underwriters. In the process, countless records and virtual traces are created that offer insight into natural phenomena, cultural phenomena and cognitive processes. A variety of approaches to the study of visual signs and meaning online have emerged across the past two decades, ranging from the close analysis of webpage layouts, to the automated analysis of aggregate data, and others in between (see e.g., KRESS; VAN LEEUWEN, 2006, CARA, 2018, MARGOLIS; PAUWELS, 2011, O'HALLORAN, 2015). Many additional methodologies wait to be developed and/or generalized; but many of those already in use are only partially described or seem inaccessible to researchers without advanced training in computer science and proprietary software. Addressing gaps like these is important for the development of semiotic methods suitable for a wide range of researchers, including digital humanists, cultural theorists, and semioticians proper.

In response, I select one such underthematized methodological focus in this paper: the use of online image search tools like Google Images² to research questions concerning human culture and cognition. My thesis and purpose in doing so is twofold: (1) to demonstrate for entry-level researchers, and for experienced researchers whose work has previously

² [google.com/imghp](https://www.google.com/imghp). As I discuss later in the paper, many researchers are already beginning to use Google Images search for carrying out empirical research, but studies that do so typically either omit semiotic theory altogether or focus their application of semiotics exclusively on social semiotics. And even the latter approaches lack systematic guides to methodology for doing so, with the notable exception of Pritchard (2020).

fallen outside the digital realm, specific ways in which systematic empirical research of visual semiotics online is already within reach, without the need for special software and/or programming skills and (2) to provide a more flexible training guide for anyone interested to apply and adapt such methods for their own ends. In the process, I provide a state-of-the-art snapshot of studies that are already using Google Images as a data collection tool, and I sketch out a step-by-step reconstruction of two of my own research projects that rely on this tool for data collection. Both cases represent ongoing research projects driven by cognitive semiotic questions—attempting to better understand and explain phenomena that arise at the semiotic intersection of anthropology, linguistics, philosophy, history, ideology, kinesiology, and embodied cognition. Both cases are also germane to the digital humanities since each necessarily entails the creation and curation of novel digital collections that may eventually be of use to other researchers.

First, it will be helpful to review the background relations and distinctions surrounding such research in order to better appreciate the ways in which qualitative research methods, like visual content analysis, map on to multimodality research, among other connections. Following this discussion, the third section of the paper provides an orientation to the cross-disciplinary literature using online image search tools for data collection and analysis. This also allows for the definition of a gradient cline of practices from the tacitly semiotic to the more overtly semiotic uses of image search tools. The two central sections of the paper then detail one case study each. In the first, I present methods used in the comparative analysis of brand mark logos to test and refine a theory of embodied semiotics involving oppositional relations. In the second, I detail methods used in the comparative analysis of images depicting the Tibetan Wheel of Life and Yama the monster of death, in order to test the embodied grounding hypothesis for the semiotic square.

I then follow up on themes that emerge from these two case studies and clarify a synthesized, step-by-step summary of methods for application to analogous research questions elsewhere. Issues of hypothesis formation, research parameters, data collection, database construction, operationalization, coding parameters, open data archiving and related issues are systematically addressed in order to further develop practices suitable for researching visual semiotics online in the context of digital humanities scholarship. In the penultimate section, I discuss issues of open data archiving, along with metadata and coding curation. The paper then closes with a summary discussion of findings and suggestions of further development.

Visual Content Analysis: background relations and distinctions

To better appreciate the relevance and situation of this methods-oriented paper, it is helpful to review a number of interrelated approaches to qualitative and mixed-methods research that intersect with theories and methods of visual semiotics and with semiotic approaches to the visual research online, with a special focus on a qualitative method known as “visual content analysis”.

First, it is helpful to consider relationships between qualitative research and semiotic inquiry. Textbook surveys of qualitative research methods often pigeonhole semiotics, mapping the study of signs and meaning primarily onto (post)structuralist frameworks premised on the identification of oppositional relations between cultural concepts (see e.g., GRBICH, 2013, p. 169-175). Some textbooks in applied semiotics also favor such approaches (e.g., HÉBERT, 2020). These approaches to semiotics are useful and valid, but they are not the sum of semiotics. General semiotics is more broadly oriented to the development of theories and methods of sign systems, meaning, and communication that address all levels of human experience and understanding. General semiotics also attempts to integrate theories and methods across disciplines, seeking valid ways of mapping between human conceptual systems and the biological dynamics from which these systems have evolved (SEBEOK, 2001; COBLEY, 2018). The theoretical infrastructure for this more general approach grows primarily from the thought of Charles S. Peirce (1839–1914) who integrated the logical study of sign systems with an inquisitive, open-ended pragmatist philosophy in which thought and action are necessarily intertwined. For Peirce, knowledge is always incomplete and human inquirers are always fallible; but, since everything is assumed to be related to everything else, however remotely, it is possible to discover incrementally the truth about things, especially when working in dialogue with various communities of inquiry.

A general approach to semiotic inquiry, then, attempts to be methodologically open (seeking to integrate instead of isolate), theoretically aware (probing critically and reflexively), logically rigorous (being tied in some way to systematic sign theory), and practically oriented (allowing new results of thought, observation, and experiment to change one’s behavior and beliefs). In fact, many approaches to qualitative inquiry are vaguely or partially semiotic without registering the connection. This is true, for example, of Grounded Theory approaches (GLASER; STRAUSS,

1967), hermeneutic methods (BERNSTEIN, 1983), phenomenological inquiry (GROENEWALD, 2004), ethnographic inquiry (WOLCOTT, 2008), content analysis (KRIPPENDORFF, 2004), and methods of triangulation (WEBB *et al.*, 2000). Indeed, such approaches often influence each other or are found to overlap substantially when compared.

Such approaches are also often independently compared to detective work, and it is here that a unifying theme emerges. A successful detective must pursue an open line of inquiry by following up on clues (many of which will be false leads) to build a case with mounting evidence until some testable conclusion can be reached. The level of inquiry that guides this process more than any other is something Peirce classifies as “abductive reasoning”, otherwise known as “guesswork”. And this is the overarching theme which situates qualitative inquiry within C. S. Peirce’s general semiotic: a point made best by Thomas Sebeok in an essay comparing Peirce to Sherlock Holmes (SEBEOK; UMIKER-SEBEOK, 1983).

With these connections in mind, the embattled distinction between quantitative and qualitative research becomes inconsequential. A blend of quantitative analysis is often useful at some level in a qualitative study, and in the case of quantitative research, the admission of qualitative guesswork (qua conceptualization, mediation, interpretation, etc.) is inescapable if the findings in question are to mean anything at all. This is just as true of visual research methods as it is of methods that focused on text and speech. Nonetheless, the quant/qual distinction remains influential in the description and classification of visual research methodologies (see, e.g., LOBINGER, 2017).

Of the many methods one might use to study phenomena in the visual domain (see e.g., MARGOLIS; PAUWELS, 2011), the method most germane to this paper is known as “visual content analysis” (VCA). Bell (2004, p. 10) defines VCA as “the explicit, quantifiable analysis of visual content as a research method”. VCA is a mode of empirical research based on patterned observations that lead to classifications of categories that can be explicitly defined (i.e., “operationalized”) between a set of images that belong to some pre-defined common class. Observed patterns are then tabulated, quantified and compared in order to test, refine and generate hypotheses relevant to one’s research questions. VCA is often used, for example, to analyze trends in magazine covers. Numerous longitudinal studies of magazine covers use VCA to discover and interpret evidence for cultural shifts (BELL, 2004; WU, 2015; O’HALLORAN, 2015, p. 402-403) by paying attention to such variables as light and colour, proxemics, cloth-

ing choices, footwear, and the interaction of text and image in order to better understand specific cultural dynamics. Depending on the scope of a given question and the size of a given dataset, vca approaches range from mechanized/algorithmic processing (O'HALLORAN, 2015; WORRING; SNOEK, 2009) to manual analysis of more manageable datasets (e.g., BELL, 2004; WU, 2015).

vca is a semiotic method, but systematic links with semiotic theory are often neglected by researchers who apply the method and by the qualitative methodologists who describe it. Notably, from a Peircean semiotic perspective, the method relies on the successful comparative observation of visual iconic signs: patterns of observed resemblances between things. According to Peirce (1903, CP 2.276–277), materially grounded icons such as those in the visual field can be distinguished from more general iconic signs by being designated “hypoicons”. These, in turn, can be further subdivided into a threefold relation between image, diagram, and metaphor: i.e., the image itself, its diagrammatic relations, and its meaning potential. For purposes of vca, diagrammatic relations are focal: the intra- and interrelationships between part–whole membership sets. Other links between semiotic theory and vca (see HUNTER, 2015) include the analysis of connotative vs. denotative meaning (following BARTHES, 1964, 1970) and the integration of Peircean semiotic to explore aspects of connotative meaning as iconic (resemblance-oriented), indexical (cause/effect-oriented), and symbolic (convention-oriented) meaning in a given image.

Visual image analysis is not the exclusive focus of vca however. Practitioners are usually just as interested in text and context, insofar as they relate to an image in question and its broader set. Because of this, vca is a semiotic enterprise in another sense as well: i.e., by attending to relationships between different styles of communication. While many researchers have begun calling such research “multimodal”, following Kress (2010), others hold that this term is too ambiguous (ZLATEV, 2019) or argue that it should be reserved as a description of the sensory modalities (sight, hearing, taste, touch, smell) as distinct from modes of communication (STAMPOULIDIS, 2019). In line with such critiques, it would be more accurate instead to refer to vca as a “polysemiotic method”. Nonetheless, the term “multimodality” is still very much in vogue and on the move (WILDFEUER, 2019); so it cannot be easily discounted or discarded.

Following the advent of personal computers and the public accessibility of images online, vca has come to be focused on predominately digitized visuals, which has led to its relevance for “multimodal digital se-

miotics” and “multimodal digital humanities” (O’HALLORAN, 2013, 2014, 2015; RAVELLI; VAN LEEUWEN, 2018). As O’Halloran (2015) points out, multimodal (i.e., polysemiotic) approaches to digital research and data curation hold much promise for the ongoing development of the digital humanities, as do semiotic applications in general. I return to this point later in the paper. For the time being, it will be helpful to further situate vca in the context of applied semiotics and image search applications online.

Applied Visual Semiotics and Google Images Search

Google Images search is already in widespread use as an empirical research tool in projects exploring questions from media/culture studies, the digital humanities, and many other fields besides. In this section I provide a brief survey of the ways in which online visual search engines like Google Images are being used for large-scale empirical analyses across an array of disciplines. I then situate such work within the domain of applied general semiotics and provide a sketch of visual content analysis as a semiotic method when using Google Images as a data collection tool.

Google consistently leads other online search engines with a market share of more than 70% (NMS, 2020), making the Google Images search tool a popular choice for browsing, searching, and researching images online. While its most common use is for ad hoc image searches and image browsing, systematic research applications are now flourishing. Even so, few published studies used the tool as a methodology prior to 2015.

In order to capture a snapshot of the methodology as it is emerging, I undertook a standard bibliographic database search to identify research studies using the key term “Google Images” and “Google Image search”. I then manually controlled for studies referencing the search engine in a non-methodological capacity. This resulted in the identification of 24 published journal articles that use the search engine as a key component of their visual research methodology. The results point to a genuinely nascent methodology, with all studies but one being published between 2014–2020. The earliest study known to adapt Google Images search for a visual research methodology is Rodriguez e Asoro (2012).

This set of 24 articles represents a dozen or so different fields, including agriculture, general biology, computer science, geography, health and medical sciences, intercultural relations, linguistics, maritime studies, marketing, neuroscience, and sociology, with many studies involving

interdisciplinary inquiry. Research questions are also diverse, including topics involving critical race studies (MANERI, 2020), border tourism (HUNTER, 2015), political cartoons (DOLATABADI; TARI, 2019), reproductive health (CHATTERJEE, 2018), otoscopic image display (CRUNDWELL *et al.*, 2015), otoscopic diagnosis (LIVINGSTONE; CHAU, 2020), midwifery (BOWDEN, SHEEHAN; FOUREUR, 2016), representation of trans surgery (MARSHALL *et al.*, 2018), burn diagnosis (PRIDGEN *et al.*, 2019), public perception of occupational therapy (WALSH, 2018), directional bias in sagittal brain representation (WISEMAN; OWEN, 2017), comparative-linguistic conceptualizations of morality (YU, WANG; HE, 2016), genetic engineering (RODRIGUEZ; ASORO, 2012), *in vitro* meat production (STEPHENS; RUIVENKAMP, 2016), spatial patterns of phenotypic traits in wild animals (LEIGHTON *et al.*, 2016), interspecific animal behaviour in the wild (MIKULA *et al.*, 2018), variation in fish coloration during breeding season (ATSUMI; KOIZUMI, 2017), anti-branding communication (KUCUK, 2015), gender arrangement in high profile social organizations (VELASQUEZ, 2016), machine learning of object categories (LIU, SHI; SHI, 2014), machine learning of agricultural categories (WSPANIALY, BROOKS; MOUSSA, 2020), machine learning of ship detection and classification (KIM *et al.*, 2018; LORENCIN *et al.*, 2019), and river bank erosion (BANDYOPADHYAY, GHOSH; DE, 2014).

As discussed above, all such studies are implicitly semiotic; though the vast majority (71%, n=17) make no explicit reference to (or application of) semiotic theory, with few even managing to frame discussions of their methodology in terms of content analysis (two exceptions being Walsh (2018) and Rodriguez e Asoro (2012)). And while the remaining seven studies attempt to address semiotic questions overtly, approaches vary. Four (BOWDEN, SHEEHAN; FOUREUR, 2016; CHATTERJEE, 2018; MANERI, 2020; STEPHENS, STEPHENS; RUIVENKAMP, 2016) apply principles of multimodal discourse analysis drawn from Social Semiotic theory (KRESS, 2010); one study applies semiotic theory drawn from both Barthes and Peirce (HUNTER, 2015); one engages in semiotic discussion with no clear application of specific theories (KUCUK, 2015); and one applies conceptual metaphor theory from cognitive semantics (YU, WANG; HE, 2016). The final instance can be counted as a cognitive semiotic theory.

These findings represent the current state-of-the-art in using Google Images search as an empirical research tool. Approaches tend to vary widely between sources, and most studies appear to be theoretically under-informed. Furthermore, as Pritchard notes, due to “Space restrictions in typical empirical publications [...] methodological accounts are often

too brief to serve as a guide” (2020, p. 297). In response to these gaps, Pritchard (2020) proposes a structured, systematic methodology for researching web images informed by “compositional, reflexive and semiotic analysis”. She refers to the approach as “combined visual analysis” (CVA). Pritchard’s CVA methodology arises from, and is focused on, a human resource management context. In this context it is “pertinent to ask how HRM can be seen via the images used in its representation and what this means for the representation of people at work.” The resulting methodology is shaped by this focus in ways that enable systematic discussion but limit broad applicability. This focus also influences Pritchard’s selection of social semiotics as a theoretical paradigm. She recommends that this specific approach to semiotic analysis should be applied as a third-stage layer of interpretation, preceded by two earlier stages of categorization and thematic analysis, each of which asks specific questions (2020, p. 299):

1. Compositional Category: What are these images of?
2. Compositional Theme: How are these images constructed?
3. Semiotic Analysis: What might these images mean?

Under each stage, Pritchard recommends that a four-part process of “Readiness”, “Recognition”, “Refinement”, and “Reflection”, should be undertaken, covering everything from data collection and categorisation to critical analysis and reflexive questioning.

Pritchard’s CVA approach is admirably systematic and may work well for those with research questions and temporal limits similar to her own. For broader research questions, with more exploratory aims, and more open-ended, longitudinal time frames, a similarly systematic guide to alternative approaches would be useful. Since it is impossible for any methodological description to account for all project variation, a comparative account is more likely to flesh out and illustrate which aspects of such an approach are more flexible and which are less-negotiable. For this reason, I discuss two empirical case studies in the following two sections. Both represent ongoing research projects, and both are focused on diagrammatic part-whole relations within and between images. Both studies are also focused on cognitive semiotic questions, asking what stylized depictions of human bodily forms can tell us about the evolution and meaning of human cognition. In the process, I identify a three stage, nine-step iterative process that allows for both open exploration of complex datasets and rigorous empirical testing:

1. Abduction
 - a. Questions & Hypotheses
 - b. Observations & Conceptualizations
 - c. Revisions
2. Deduction
 - a. Data Parameters & Operationalization
 - b. Data Collection & Input
 - c. Revisions
3. Induction
 - a. Database Development & Coding
 - b. Database Analysis & Interpretation
 - c. Revisions

Instead of somehow relegating semiotics to a third phase, this model is semiotic through and through, and each stage is also interwoven, virtually and actually, with every other stage. For this reason, the Peircean Abduction-Induction-Deduction schema is particularly fitting. According to Peirce, whether they are involved in logical abstraction or empirical inquiry the triadic categories are always interdependent in spite of their irreducibility. The repetition of “Revisions” at each level is as much an acknowledgement of the necessary interweaving of the three modes of inquiry as it is an acknowledgement of the Fallibilist nature of inquiry in general: a process requiring ongoing interpretation and reinterpretation in relation to processes of infinite semiosis.

Case Study 1: comparative X-Posed Brand Marks

Primary Phase: Abduction. The first case study grows out of broader interests in the exploration of open-ended, interrelated questions surrounding embodied semiotics, symmetry theory, and the nature of oppositional relations in human cognition. Among the many miscellaneous notes related to these topics that I continue to file away informally in a general database³ for possible future exploration, I began to incidentally amass a small but curious collection of brand-mark logos featuring sym-

³ I personally tend to use Microsoft OneNote for this purpose until a collection becomes unmanageable—a good sign that a more formal, systematic research project is called for.

metrically (or quasi-symmetrically) designed human figures doing full body X-poses, also known as “spread-eagle” formation. I didn’t know what to make of these tokens but found them intriguing and potentially worth looking into further down the road. As the collection grew incrementally from 2011–2014, based on chance discoveries, a number of questions began to emerge regarding variations and contextual associations. Some tokens were sub-divided at the waist; others were subdivided along the sagittal plane; others were fully integrated; some featured props or internal design elements (see Figure 1).

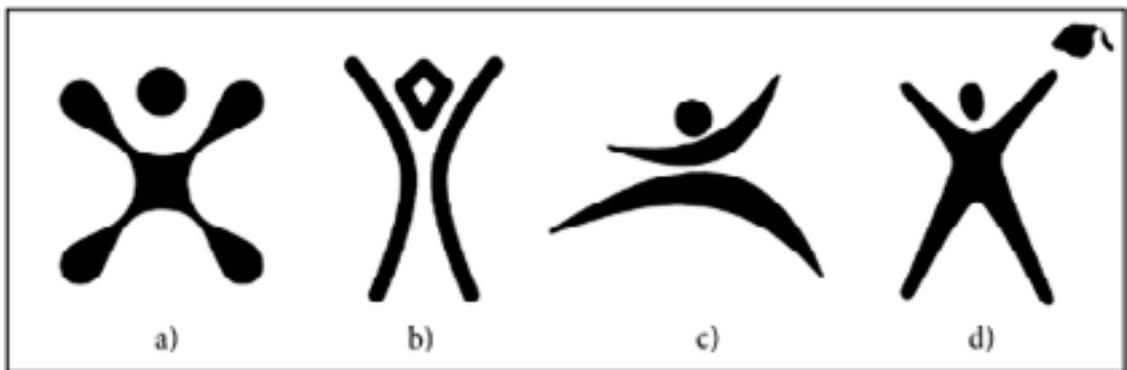


Figure 1. X-posed logo exemplars: (a) Cingular Wireless logo (2000-2008: United States); (b) Burning Man logo (Nevada, USA); (c) Jump4Joy logo (Netherlands); (d) Avid logo (USA).

Would a larger sample reveal tendencies in favour of one or another of these design elements? Would these preferences map onto contextual semiospheres in some way? Then there were questions of the semiospheres themselves: What kinds of businesses used these logos? Did the logos have dominant thematic or semantic fields? Were certain kinds of corporations or services represented more than others? I decided that a more formal study featuring a larger sample set of these designs was in order. The research tool of choice? Google Images. The goal? To build a database of 100 to 200 exemplars suitable for statistical cross-tabulation for generating, testing, and refining hypotheses related to these representations of human embodied form. The database in question remains open, but most data under discussion in this section I added during an intensive two-year period of research, from 2015–2016.⁴

Secondary Phase: Deduction. Prior to beginning the Google Image search process—or, rather, part and parcel with the early process of doing so—it was necessary to define what should and should not constitute a suitable exemplar; otherwise, the objects under investigation would have no grounds for valid comparison and any resulting studies would lack ori-

⁴ During a crucial four-month stretch of data collection in 2016, my efforts were aided by Sean Murray, a MA research assistant funded by the Ryerson-York Graduate Program in Communication and Culture.

entation and focus. The deductive process in question is often discussed as “operationalization”. As I describe in a preliminary write-up on findings from the database (PELKEY, 2017b, p. 67), two key criteria emerged for operational screening of individual cases in this regard: One functional, the other structural:

1. Functional Criterion for Inclusion: Actual Usage. *The logo in question must be in actual use by a functioning, registered organization (whether for-profit or not-for-profit, presently or historically), as either a product brand mark or corporate brand mark.*
2. Structural Criterion for Inclusion: Singular Posture Type. *The logo in question must include a solitary (one and only one), spread-eagle human form, including identifiable representations of head, arms and legs, with arms raised and legs parted at obtuse angles relative to the vertical axis of the torso.*

These criteria allowed for the inclusion a range of valid cases while setting clear screening limits for excluding non-valid cases from the database. The first criterion excluded the endless supply of mere stock graphics or concept designs available online while helping to ensure that each logo being compared would be situated within its own semiosphere of associations, allowing for robust comparisons of actual semiotic variables. The second criterion defines the posture itself and distinguishes it from related postures or multiple performances of the same posture within a single logo.

Data collection proceeded hand-in-hand with the definition of these criteria during its early stages, requiring the exclusion of some data that had initially been included. Once operational screening criteria were clearly established, Google Image searches also became increasingly systematic. Early searches used linguistic keyword searches such as “X logo”, “happiness logo”, “pain logo”, following up on semantic domain clues generated from collected data. As particular professions began to surface, these were added to keyword searches, including “extreme sports logo” and “chiropractor logo”. Since the goal of the database was to identify and as many valid exemplars as possible, there was no reason to artificially define search strings in advance. Instead, new thematic findings generated new potential search strings. As the database grew, keyword searches transitioned to reverse image searches, allowing Google’s image matching algorithm to search for images with visually similar features. Carrying out searches in a private browsing mode with cookies disabled proved

to be important for ensuring new search criteria were not being filtered through earlier search criteria. This process proceeded until no new valid images were being returned in search results. Using these methods, the database grew to include more than 200 exemplars. Metadata collection proceeded hand-in-hand with the addition of new cases to the database, with all information being added to a single spreadsheet, inclusive of with the following potential variables:

1. Logo Code: A unique abstract identifier for each case that matches the filename of its namesake
2. Logo Name: A shorthand term or phrase to identify the represented company or product
3. Domain 1: An upper-level thematic domain for classifying the logo relative to others
4. Domain 2: A thematic subdomain for classifying the logo relative to others
5. Primary URL: The original, principal access URL situating the logo in its contextual website
6. Date: Specifying when the logo was added to the database
7. Status of Use: Whether or not the logo has been discontinued in spite of historical usage
8. About Us URL: Web link to the About Us page describing the company or product in question
9. Country of Origin: Nation state in which the company using the logo is located
10. External URL: Any auxiliary website with relevant information about the company or product
11. Slogan/Tagline 1: Brief phrase used in marketing the product or company online
12. Slogan/Tagline 2: Secondary marketing slogan (if applicable)
13. Executive Summary: Longer statement of company/product values and purpose

14. Logo Name: Proper name used by company to identify and discuss its logo (if applicable)
15. URL About Logo: Web link to page discussing logo directly (if applicable)
16. Notes: Any additional points of clarification or context useful for interpretation

Tertiary Phase: Induction. The third interrelated phase of inquiry involves actual data analysis: the testing of hypotheses and interpretation of results. This requires the selection of codeable variables from existing data and metadata, and further classification of figures based on observable features in order to generate cross-tabulations that could be reported as general statistics and trends or calculated to discover contingency coefficients. These results, in turn, serve as evidence for making higher-order hypothesis related to broader meanings and more general relationships with findings elsewhere. For this phase of the X-Posed Brand Mark project, the process also involved working in dialogue with a research team to establish perceptual dimensions of contrast in the data and to begin to test cross-coder reliability.

In practice, this meant that each logo in the database needed to be analyzed relative to a set of select variables in a process that can be classified as visual content analysis (vca: see earlier discussion in Section 2). For purposes of this case study vca variables could be textual, thematic, conceptual, morphological, and/or anatomically symmetrical. Fifteen dimensions of contrast emerged through processes of team discussion and open-coding, including many of the variables introduced above along with seven coding variables related to the anatomical planes and symmetrical or asymmetrical dimensions of contrast drawn from human physiology and theories of plane pattern analysis (WASHBURN; CROWE, 1988, 2004; MARSDEN; THOMAS, 2013). Although much more analysis waits to be carried out using this database, preliminary results (pending further data collection and coding) suggest a number of remarkable findings that lend themselves well to interpretation against a broader backdrop of connections in a larger study, as reported in Pelkey (2017b, p. 63–83). Morphologically speaking, a clear typology of patterns emerged, emphasizing the organizing influence of the anatomical planes (as illustrated in Figure 1).

Thematically speaking, at-risk demographics (e.g., children, youth, refugees, the poor, the elderly) and risky behaviours (e.g., gambling addiction, extreme sports, electrical services, insurance providers) both

emerged as salient themes motivating the use of spread-eagle logos. Extreme experiences of both pain and pleasure were also found to be well-represented, with health-and-wellness-oriented companies being the most frequently represented and with chiropractic emerging as the best represented occupation in the database. A full 30% (n=60) of logos represent chiropractic clinics, where people with extreme pain go for extreme procedures for pain relief. Clues such as these all provide further evidence of the posture's associations with extremes and reversals. Such findings also provide supporting evidence for better understanding the meanings and origins of this full body posture, along with the ways its memory traces and conceptual mappings serve to inform human cognition.

Case Study 2: comparative Yama Bhavachakra Mandalas

Primary Phase: Abduction. The second case study under consideration here grows out of the same network of research questions surrounding the semiotics of embodied patterns as the first. I refer to this ongoing line of inquiry as “embodied pattern grammar”; but, instead of looking to contemporary brand marks found primarily in western corporate culture for clues, however, this case study shifts the focus to a more historically oriented visual symbol system or model that emerges in a decidedly non-western context: Tibetan Buddhism. The model in question is a sacred didactic mandala or Bhavachakra featuring the Wheel of Rebirth or *Samsara* being supported by the monster of impermanence, also known as Yama the god of death. I will forego an in-depth description of the model here for the sake of methodological focus; but suffice it to say that these depictions each consist of several indispensable component parts that are organized recursively according to diagram-internal and cultural-internal logical relations. The two most prominent features of the diagram are the monster Yama and the wheel of life itself. My interest in this diagram type, beyond its captivating visual appeal, is the possibility that its internal organizational logic might map onto the logic of the semiotic square of A. J. Greimas (1987). The weight of this possibility in turn would be augmented by the embodied relations of the four-limbed Yama who supports it; and this orientation might provide evidence for testing my hypothesis that the semiotic square is itself an embodied diagram (PELKEY, 2017a).



Figure 2. Yama Bhavachakra exemplars. Left: painted thanka, c.1900, Tibet, via Royal Ontario Museum;⁵ Middle: embroidered thanka, c.1800, Eastern Tibet, via Wikimedia Commons;⁶ (right) painted thanka (c.1800, Mongolia), via Rubin Museum of Art.⁷

With this hypothesis in mind, I began to wonder whether or not different tokens of this model were robustly consistent in terms of their internal organization between part-whole relations or, alternatively, whether they might be marked by substantial variation. Did organizational schemes differ from region to region or through time? Was variation between diagrams patterned and regular with logical progressions or were differences relatively random? Were there dominant patterns and recessive patterns? The only systematic, historical study of the model (TEISER, 2006) helped orient myself to these questions in some ways but provided few systematic answers. And looming largest of all was the question of the wheel analogy itself. Since a wheel can rotate 360 degrees, one might expect based on the analogy alone that the wheel's internal relationships would be more randomly ordered than consistently ordered. My guess, though, was that this hypothesis would be proven wrong due to the embodied constraints of the monster Yama who provides an organizing frame of reference for the diagram and its observers. In short, the hypothesis was complex, and I had more questions than answers. Indeed, finding answers and testing these hypotheses is still an ongoing process at the time of writing, but I have already taken the next steps.

Secondary Phase: Deduction. In order to begin testing these hypotheses and interpreting my findings, I needed to define parameters. I would also need to collect a preliminary dataset with sample size and scope sufficient to make valid general claims. For the initial phase of the study, my

5 himalayanart.org/items/77565.

6 commons.wikimedia.org/wiki/File:Wheel_of_Existence.jpg.

7 himalayanart.org/items/78.

goal was to collect 100 unique cases using Google Images search in order to reach a reasonable threshold suitable for statistical inference. To get started, I once again began with keyword searches and with the handful of images that I had collected in an ad hoc manner (while focused on other questions) across the span of several years. It soon became necessary to define my operational parameters for what would and would not constitute a valid case.

As mentioned above, my first two criteria were the monster and the wheel. Some version of Yama the god of death needed to be present, with two hands visibly protruding from the top half of a wheel diagram, and two feet visible protruding from the bottom half. The wheel, in turn, needed to be segmented in some visible way to, at least, indicate the various realms into which one might be reborn. While there is much more to most Tibetan wheel of life diagrams than these elements, and while many of these other elements are also important for answering the questions and testing the hypotheses introduced above, these three emerged as the minimal qualifying components needed for operationalizing a valid case and screening for invalid cases.

Once I had amassed a core starter set of 25 cases from both earlier ad hoc collection and new keyword searches (i.e., Bhavachakra, Tibetan Wheel of Life, Wheel of Life), I began to rely solely on reverse image searches for discovering more cases. In this way, I quickly found that my most fruitful results stemmed from collecting further cases on webpages linked from the Google Images search page. In order to document this process of data collection off-site, I found it necessary to a Google Images hit tracking column to my metadata spreadsheet. In this way, I was able to record, for example, that reverse image searches involving image case 3/25 resulted in the collection of 32 new images: 9 from the main Google Images page (3.1–3.9), 20 from a website⁸ linked via image case 3.7 (3.7.1–3.7.20), and three from a website⁹ linked via image case 3.8 (3.8.1–3.8.3). These records help provide an important layer of transparency useful for validating or scrutinizing the data collection process. I included six other layers of metadata as well for documenting and contextualizing images—many of which may themselves prove to be useful as coding variables in future analyses:

8 himalayanart.org.

9 traditionalartofnepal.com.

1. ID#: Data case identifier
2. Webpage layer tracking: layer tracking coding system to identify embedded page searches.
3. Source: Organization hosting the image online
4. Web link: The source url for locating the image online
5. Country: The nation state of origin (i.e., where the image was created vs. curated).
6. Era: The year or approximate era in which the image was originally produced (if known)
7. Medium: The material and medium used for producing the original artwork
8. Notes: Miscellaneous clarifications including curatorial location of image (if applicable)

Tertiary Phase: Induction. Once I had managed to collect at least 100 unique image cases (the current database stands at 103), it was possible to begin analysing and coding the patterned relationships within and between these images. This was done in preparation for, and co-requisite with, analysis. In other words, as I started to analyze the data, I slowly found that I needed to find answers to a number of new questions, prominently including the following:

1. Realms: how many identifiable realms of *Samsara* are featured in each wheel?
2. Spokes: how many delineated spokes separate the realms of *Samsara*?
3. Samsara Pattern: what are the patterned relations of actor types in each realm?
4. Unmarked Position: which actor type is depicted in the upper-left section of the wheel?
5. Correlative Position: which actor type is depicted in the lower-left section of the wheel?
6. Karma Pattern: if the hub of the wheel depicts a karma distinction, is the subdivision vertical or horizontal?

In each case, the answers to these questions are capable of enumeration (being turned into a numerical value in a coding spreadsheet) that can then be cross-tabulated with other variables in order to report basic statistics, or to calculate contingency coefficients useful for making strong claims about general tendencies and relationships in the data. Once again, methods of vca and grounded theory coincide substantially with this level of analysis. The semiotician is simply able to draw on a more varied array of relationships, asking and answering questions more consequential in scope by situating the methodology in a more general, pragmatist approach to semiotic.

In the process of doing so for this particular case study, I made a number of remarkable, if tentative, discoveries and interpretations in need of further testing. Since these findings are currently unpublished, since their description would require much more contextualization than space allows here, and since the purpose of this paper is methods-oriented, let me simply say that the induction phase of this research has resulted in interpretive layers that suggest these diagrams are a valid source of evidence for affirming the embodied grounding hypothesis for the semiotic square due to their patterned organization.

Generalizations for Visual Semiotics Research Online

These two methodological sketches of visual semiotic research online using Google Images are somewhat artificially subdivided into three discrete phases each. This segmentation is helpful for heuristic purposes, but it is equally helpful to consider how, in actuality, there is a constant movement between abduction, induction, and deduction in processes of inquiry, as Peirce describes (1903, EP2, p. 208-225). Even though this movement is difficult to codify and outline, it is a familiar part of our everyday experience. Peirce notes, for example, that “the whole fabric of our knowledge is one matted felt of pure hypothesis confirmed and refined by induction. Not the smallest advance can be made in knowledge beyond the stage of vacant staring, without making an abduction at every step” (1901, HP 2, p.900). This is equally true in carrying out formal research.

When moving into the analysis of the Bhavachakra diagrams, for instance, I immediately needed to generate new hypotheses upon discovering differences in the number of Samsara realms between images and variant ways of marking these differences. This, in turn, required new deductive strategies for operationalizing these distinctions. But in order

to settle on such definitions, I needed to resort to further abduction and induction as well. In deciding, for example, that the presence of bows and arrows could differentiate the titan realm from the human realm, I had to first guess that this might be the case (abduction), then assume it to be the case (deduction), and then test to see whether or not it was actually the case (induction). Exceptions to the rule could then help refine the original hypothesis. These dynamics help explain the addition of “Revisions” in the “three stage, nine-step iterative process” outlined in Section 3 above.

Returning to questions of digital methodology using Google Images search, it is also important to note that both case studies moved beyond front page “top slice” searches relying solely on linguistic search strings. Instead of assuming that the first *x* number of images in a search return would provide a valid sample set (something that many other studies reviewed above recommend), these two research projects called for the collection of a much larger sample. Because of this, multiple keyword searches were complemented by reverse image searches and off-site data collection following successively embedded links. While a top slice approach may be adequate for projects with basic, pre-defined research questions, I recommend the more complex strategies introduced above when research questions are more multifaceted and exploratory, requiring a larger comparative database. Both of these methodological case studies are also open-ended in the sense that they are both ongoing projects. Projects that require a more bounded time-frame, with more manageable research questions, may benefit more from a top-slice approach that hews closer to Pritchard’s thematic/perceptual analyses of human resource management depictions through a Google Images top-slice approach (PRITCHARD, 2020).

This is not to suggest that rigid parameters do not apply to more open-ended studies. It is important to note, in this regard, that in each of the two studies introduced above, new constraints were often necessary in addition to those assumed at the outset. In spite of being formulated prior to commencing with formal data collection, for instance, operational screening criteria for data case inclusion (or exclusion) needed to be refined during the process of data collection. This refinement resulted in the disqualification of some items and in the creation of alternative databases for future projects. In the X-posed brand mark study, for example, cases featuring two or more X-posed figures were ruled out early in the data collection process as a way of limiting the scope of the database and enhancing the validity of comparisons between cases.

Data Archiving and Semiotic Stewardship

Semiotic research is a fundamentally comparative enterprise (SONESSON, 2016). Any discussion of semiotic research methods is then an attempt to enhance our comparative abilities in identifying and interpreting the meanings of semiotic systems. The enhancement of comparative *capacity*, comparative *validity*, and comparative *practice* are, among others, all worth considering.

Enhancements to our comparative capacity are made possible by online search tools such as Google Images. Enhancements to the validity of our comparisons are made possible by critical evaluation sharpened in dialogue with specific theories, methods, and communities of inquiry. Enhancements to our comparative practices are enabled by the careful curation of collected visual data and the creation and input of metadata fields for tracking and managing this information. Each of these enhancements are vital for the empirical research processes described above. But what about comparative *availability*? How do we ensure that others are able to test our results and participate in our comparative research? Due to the constantly shifting nature of hyperlinks and webpages, it is frequently impossible to reproduce even the most basic and carefully documented dataset.

Fortunately, then, it is now possible to make one's datasets (visual or otherwise) widely available using public upload repositories and registered hyperlinks or digital object identifiers (DOI) thanks to free online hosting services such as Zenodo.¹⁰ Zenodo is an Open Science initiative founded by the European Organization for Nuclear Research (CERN) that is open to dataset contributions and the creation of data research communities from all disciplines. Emerging guidelines, such as the open access "FAIR Principles", help ensure that data and metadata submitted to such open repositories is "Findable, Accessible, Interoperable, and Reusable" (WILKINSON *et al.*, 2016). Once the two research projects discussed above conclude, the visual data, metadata, database, and coding sheets will be added to their own linked archive on Zenodo or a comparable site.

Given the vagaries and mutability of copyright laws from region to region and year to year, the ways in which, or the extent to which, visual data should itself be added to open repositories are themselves open questions. As various other researchers have described, it is currently impossible to suggest universal guidelines for publicly archiving visual data

¹⁰ zenodo.org.

collected from the internet (BOJE; SMITH, 2010; PRITCHARD, 2020). Instead, after consulting one's local guidelines, it is more useful to appeal to fair dealing or fair use guidelines for non-profit, educational purposes, in addition to practicing careful, accurate, and detailed attribution of image sources in the archived metadata. As more and more visual research data is curated in these ways, digital humanists, visual anthropologists, and other visual semioticians will be able to establish more detailed guidelines and practices that eventually coalesce into well established precedent. Until then, thoughtful experimentation seems warranted for the sake of semiotic stewardship.

Conclusion

Digital humanities scholarship has much to gain from enriched integration with general semiotic theory and methodologies such as those discussed in this paper. Conversely, approaches to general semiotic research have much to gain from closer dialogue with digital humanities scholarship and practice. Both approaches need to develop more flexible, robust methodological guidelines for carrying out visual research online. That has been my goal in this paper, with a special focus on the use of Google Images as a research tool. I have also attempted to describe ways in which such research could move beyond applied semiotics to contribute reflexively to the development of semiotic theory and methodology.

The sheer scope, variety, and availability of visual data that online image search platforms offer is unprecedented. One affordance these tools introduce is increased access to variations in part-whole relationships between validly classified image types. In the words of Atsumi e Koizumi (2017, p. 567), "Web image analyses are still preliminarily with many limitations, but could be promising for investigating variations in visible traits". Paying attention to such variations, on the other hand, requires focused training, filtered through theoretical paradigms, and applied through suitable methodologies. Such layers of attention are known in Peircean semiotics as "interpretants", that which one is prepared to interpret when noticing some representamen-object relation. With robust interpretants, one can not only identify and classify variation patterns, one can also use those findings to contribute to semiotic theory.

In this paper, I presented two case studies that demonstrate this potential. In both cases, paying close attention to variations between individual images relative to specialized interpretants enabled the discovery of new insights useful for answering questions in cognitive semiotics. The first study illustrated the comparative analysis of variations in X-posed

brand mark logos, the second study illustrated the comparative analysis of variations between Tibetan wheel of life diagrams. Both studies are ongoing, and both have already resulted in findings that help test and refine theories of embodied semiotics involving logical relations of opposition and correlation.

The approach involves a sustained, multi-layered analysis of systematic relations or “diagrammatic iconicity” (Nöth, 2008) developing standard practices in vca. As such, it serves as a complementary approach to the systematic methodology recommended by Pritchard (2020) for Google Images research geared toward perceptual/thematic studies involving shorter time constraints and top-slice approaches to data collection. Another way in which my approach differs from Pritchard’s is in my assumption that the methodology should be mapped onto semiotic theory and praxis from start to finish (instead of relegating semiotics to a third, interpretive step). This was accomplished by situating the methodology within Peirce’s three modes of inquiry: abduction, deduction, and induction. I then went on to highlight the many ways in which these phases overlap, since processes of guesswork, hypothesis refinement, and interpretation are apparent in every phase. In other words, each phase of the research constitutes a kind of microcosm of the whole, in spite of being distinct, both chronologically, conceptually, and methodologically from each of the other two phases. Ultimately, then, the approach is thoroughly semiotic in at least three senses, including experiential, formal, and practical means:

1. Experiential: by blending of sensory modes and communication channels
2. Formal: by attending to hypoiconic manifestations with a focus on diagrammatic relations, mapping social semiotic questions onto analyses of icon-index-symbol as manifestations of connotative meaning vs. denotative meaning
3. Practical: by undertaking iterative detective-style guesswork in which hypotheses are adopted, tested and refined toward interpretive ends.

The fact that this approach can also be used to contribute to the development of semiotic theory and methodology makes it semiotic in a fourth way as well, but this need not be the aim of every researcher who sets out to discover answers to visually oriented questions by critically sifting through the abundance of patterned variation made available by online image search engines. Semiotics has much to offer and expects nothing in return.

References

- ATSUMI, Keisuke; KOIZUMI, Itsuro. Web image search revealed large-scale variations in breeding season and nuptial coloration in a mutually ornamented fish, *Tribolodon hakonensis*. *Ecological Research*, v. 32, n. 4, p. 567–578, 2017. DOI: [10.1007/s11284-017-1466-z](https://doi.org/10.1007/s11284-017-1466-z).
- BANDYOPADHYAY, SHREYA; Ghosh, Kapil; DE, Sunil Kumar. A proposed method of bank erosion vulnerability zonation and its application on the River Haora, Tripura, India. *Geomorphology*, v. 224, p. 111–121, 2014. DOI: [10.1016/j.geomorph.2014.07.018](https://doi.org/10.1016/j.geomorph.2014.07.018).
- BARTHES, Roland. *Elements of semiology*. New York: Hill and Wang, 1964.
_____. *Mythologies*. New York: Hill and Wang, 1972.
- BERNSTEIN, Richard J. *Beyond objectivism and relativism: Science, hermeneutics and praxis*. Philadelphia: University of Pennsylvania Press, 1983.
- BELL, Philip. Content analysis of visual images. In: VAN LEEUWEN, Theo; JEWITT, Carey (eds.). *The handbook of visual analysis*. London: SAGE, 2004. p. 10–34. DOI: [10.4135/9780857020062.n2](https://doi.org/10.4135/9780857020062.n2).
- BOJE, D.; SMITH, R. Re-storying and visualizing the changing entrepreneurial identities of Bill Gates and Richard Branson. *Culture and organization*, v. 16, 2010, p. 307–331.
- BOWDEN, Calida; SHEEHAN, Athena; FOUREUR, Maralyn. Birth room images: What they tell us about childbirth. A discourse analysis of birth rooms in developed countries. *Midwifery*, v. 35, p. 71–77, 2016. DOI: [10.1016/j.midw.2016.02.003](https://doi.org/10.1016/j.midw.2016.02.003).
- CARA, Mariane. The Semiotic layers of Instagram: Visual tropes and brand meaning. *American Journal of Semiotics*, v. 34, n. 3–4, p. 331–352, 2018. DOI: doi.org/10.5840/ajs201931146.
- CHATTERJEE, Karishma. What can we learn about the female condom online? An analysis of visual representations of the female condom on the Internet. *Atlantic Journal of Communication*, v. 26, n. 3, p. 149–163, 2018. DOI: [10.1080/15456870.2018.1474217](https://doi.org/10.1080/15456870.2018.1474217).
- COBLEY, Paul. *Cultural implications of biosemiotics*. Cham: Springer, 2016.
- CRUNDWELL, G.; HARMER, J.; MALTBY, M.; MILLS, T.; NEUMANN, C.; WALSH, L.; BAGULEY, D. Images of otoscopy: rate and extent of non-compliance with good practice standards. *Journal of Laryngology & Otology*, v. 129, n. 1, p. 27–31, 2015. DOI: [10.1017/S0022215114003004](https://doi.org/10.1017/S0022215114003004).

DOLATABADI, Hadi; TARI, Zeinab Ghasemi. A visual framing analysis of French and us political cartoons on Trump's withdrawal from the Iran nuclear deal. *Journal of World Sociopolitical Studies*, Teheran, v. 3, n. 3, p. 605-647, 2019. DOI: [dx.doi.org.ezproxy.lib.ryerson.ca/10.22059/WSPS.2020.289255.III9](https://doi.org/10.22059/WSPS.2020.289255.III9).

GLASER, Barney G.; STRAUSS, Anselm L. *Discovery of grounded theory: strategies for qualitative research*. New Brunswick, N.J.: Aldine, 1967.

GROENEWALD, Thomas. A phenomenological research design illustrated. *International Journal of Qualitative Methods*, v. 3, n. 1, p. 42-55, 2004. DOI: [10.1177/160940690400300104](https://doi.org/10.1177/160940690400300104).

GRBICH, Carol. *Qualitative data analysis: an introduction*, 2nd ed. London: Thousand Oaks, 2012.

GREIMAS, A. J. *On meaning: selected writings in semiotic theory*. Minneapolis, MN: University of Minnesota Press, 1987.

HÉBERT, Louis. *An introduction to applied semiotics: tools for text and image analysis*. New York: Routledge, 2020.

HUNTER, William Cannon. The visual representation of border tourism: Demilitarized zone (DMZ) and Dokdo in South Korea. *International Journal of Tourism Research*, v. 17, n. 2, p. 151-160, 2015. DOI: [10.1002/jtr.1973](https://doi.org/10.1002/jtr.1973).

HUSSAIN, Aabid; GUL, Sumeer; SHAH, Tariq Ahmad; SHUEB, Sheikh. Retrieval effectiveness of image search engines. *The Electronic Library*, v. 37, n. 1, p. 173-184, 2019. DOI: [10.1108/EL-07-2018-0142](https://doi.org/10.1108/EL-07-2018-0142).

KIM, Kwanghyun; HONG, Sungjun; CHOI, Baehoon; KIM, Euntai. Probabilistic ship detection and classification using deep learning. *Applied Sciences*, Basel, v. 8, n. 6, 2018. DOI: [dx.doi.org.ezproxy.lib.ryerson.ca/10.3390/app8060936](https://doi.org/10.3390/app8060936).

KRESS, Gunther R. *Multimodality: a social semiotic approach to contemporary communication*. London: Taylor & Francis, 2010.

KRESS, Gunther R.; VAN LEEUWEN, Theo. *Reading images: the grammar of visual design*. 2nd ed. London: Routledge, 2006.

KRIPPENDORFF, Klaus. *Content analysis: an introduction to its methodology*. London: Sage, 2004.

KUCUK, S. Umit. A semiotic analysis of consumer-generated antibranding. *Marketing Theory*, v. 15, n. 2, p. 243-264, 2015. DOI: [10.1177/1470593114540677](https://doi.org/10.1177/1470593114540677).

LEIGHTON, Gabriella R. M.; HUGO, Pierre S.; ROULIN, Alexandre; AMAR, Arjun. Just Google it: assessing the use of Google Images to describe geographical variation in visible traits of organisms. *Methods in Ecology and Evolution*, v. 7, n. 9, p. 1060-1070, 2016. DOI: [10.1111/2041-210X.12562](https://doi.org/10.1111/2041-210X.12562).

LIU, Xi; SHI, Zhi-Ping; SHI, Zhong-Zhi. A co-boost framework for learning object categories from Google Images with 1st and 2nd order features. *The Visual Computer*, v. 30, n. 1, p. 5-17, 2014. DOI: [10.1007/s00371-012-0772-2](https://doi.org/10.1007/s00371-012-0772-2).

LIVINGSTONE, Devon; CHAU, Justin. Otoscopic diagnosis using computer vision: An automated machine learning approach. *The Laryngoscope*, v. 130, n. 6, p. 1408-1413, 2020. DOI: [10.1002/lary.28292](https://doi.org/10.1002/lary.28292).

LOBINGER, Katharina. Visual research methods. In: *The International Encyclopedia of Communication Research Methods*. American Cancer Society, p. 1-10, 2017. DOI: [10.1002/9781118901731.iecrm0265](https://doi.org/10.1002/9781118901731.iecrm0265).

LORENCIN, Ivan; ANDELIC, Nikola; MRZLJAK, Vedran; CAR, Zlatan. Marine objects recognition using convolutional neural networks. *Nase More*, Dubrovnik, v. 66, n. 3, p. 112-119, 2019. DOI: dx.doi.org.ezproxy.lib.ryerson.ca/10.17818/NM/2019/3.3.

MANERI, Marcello. Breaking the race taboo in a besieged Europe: How photographs of the “refugee crisis” reproduce racialized hierarchy. *Ethnic and Racial Studies*, p. 1-17, 2020. DOI: [10.1080/01419870.2020.1723672](https://doi.org/10.1080/01419870.2020.1723672).

MARGOLIS, Eric; PAUWELS, Luc (eds.). *The SAGE handbook of visual research methods*. London: SAGE, 2011.

MARSDEN, Jamie; THOMAS, Briony. Brand values: Exploring the associations of symmetry within financial brand marks. *Design Management Journal*, v. 8, n. 1, p. 62-71, 2013. DOI: dx.doi.org/10.1111/dmj.12004.

MARSHALL, Zack; BRUNGER, Fern; WELCH, Vivian; ASGHARI, Shabnam; KAPOSY, Chris. Open availability of patient medical photographs in google images search results: Cross-sectional study of transgender research. *Journal of Medical Internet Research*, v. 20, n. 2, p. e70, 2018. DOI: [10.2196/jmir.8787](https://doi.org/10.2196/jmir.8787).

MIKULA, Peter; HADRAVA, Jirí; ALBRECHT, Tomáš; TRYJANOWSKI, Piotr. Large-scale assessment of commensalistic-mutualistic associations between African birds and herbivorous mammals using internet photos. *PeerJ*, 2018. DOI: dx.doi.org.ezproxy.lib.ryerson.ca/10.7717/peerj.4520.

NET MARKET SHARE (NMS), 2020. Search engine market share. *Net MarketShare*. netmarketshare.com/search-engine-market-share.

NÖTH, Winfried. Semiotic foundations of natural linguistics and diagrammatic iconicity. In: WILLEMS, Klaas; DE CUYPERE, Ludovic (org.). *Naturalness and iconicity in language*. Amsterdam: Benjamins, 2008. p. 73-100.

O'HALLORAN, Kay. Multimodal digital humanities, In: TRIFONAS, Peter Pericles (ed.), *International handbook of semiotics*, v. 1. Cham: Springer, 2015. p. 389-415.

O'HALLORAN, Kay; CHUA, Alvin; PODLASOV, Alexey. The role of images in social media analytics: A multimodal digital humanities approach. In: *Visual communication*. Berlin: de Gruyter, 2014. p. 565-588.

O'HALLORAN, Kay L.; MARISSA, K. L. E.; PODLASOV, Alexey; TAN, Sabine. Multimodal digital semiotics: the interaction of language with other resources. *Text & Talk*, v. 33, n. 4-5, p. 665-690, 2013. DOI: [10.1515/text-2013-0030](https://doi.org/10.1515/text-2013-0030).

PEIRCE, Charles S. *The collected papers of Charles Sanders Peirce*, v. 1-6, HARTSHORNE, C.; WEISS, P. (eds.) (1931-1935); v. 7-8, BURKS, A. (ed.) (1958). Cambridge, MA: Harvard University Press. (Cited as CP followed by volume number and section number.)

_____. *The essential Peirce*, v. 2, PEIRCE EDITION PROJECT (ed.). Bloomington, IN: Indiana University Press, 1998. (Cited as EP 2 followed by page number.)

_____. *Historical perspectives on Peirce's logic of science: a history of science*. EISELE, Carolyn (ed.), 2 vols. Berlin: Mouton, 1985. (Cited as HP followed by volume and page number.)

PELKEY, Jamin. Greimas embodied: How kinesthetic opposition grounds the semiotic square. *Semiotica*, n. 214, p. 277-305, 2017a.

_____. *The semiotics of X: Chiasmus, cognition, and extreme body memory*. London: Bloomsbury Academic, 2017b.

PRIDGEN, B. C.; JOPLING, J. K.; SHECKTER, C. C.; DESPO, O.; YEUNG, S.; KARANAS, Y.; LI, F.; MILSTEIN, A. S. 33 Automated burn assessment using deep learning and computer vision. *Journal of Burn Care & Research*, v. 40, n. Suppl. 1, p. S25-S26, 2019. DOI: [10.1093/jbcr/irz013.037](https://doi.org/10.1093/jbcr/irz013.037).

PRITCHARD, Katrina. Examining web images: a combined visual analysis (CVA) Approach. *European Management Review*, v. 17, n. 1, p. 297-310, 2020. DOI: [10.1111/emre.12376](https://doi.org/10.1111/emre.12376).

RODRIGUEZ, Lulu; ASORO, Ruby Lynn. Visual representations of genetic engineering and genetically modified organisms in the online media. *Visual Communication Quarterly*, v. 19, n. 4, p. 232-245, 2012. DOI: [10.1080/15551393.2012.735585](https://doi.org/10.1080/15551393.2012.735585).

RAVELLI, Louise J.; VAN LEEUWEN, Theo. Modality in the digital age. *Visual Communication*, v. 17, n. 3, p. 277-297, 2018. DOI: [10.1177/1470357218764436](https://doi.org/10.1177/1470357218764436).

SEBEOK, Thomas A. *Global semiotics*. Bloomington, IN: Indiana University Press, 2001.

SEBEOK, Thomas A.; UMIKER-SEBEOK, Donna Jean. "You know my method": a juxtaposition of Charles S. Peirce and Sherlock Holmes. In: ECO, Umberto, SEBEOK, Thomas A. (org.). *The sign of three*. Bloomington, IN: Indiana University Press, 1983. p. 11-54.

SONESSON, Göran. The phenomenological semiotics of iconicity and pictoriality – Including some replies to my critics. *Language and Semiotic Studies*, v. 2, n. 2, p. 1-73, 2016.

SPENCE, Des. Imagine all the people. *BMJ: British Medical Journal*, v. 346, n. 7894, p. 39-39, 2013.

STAMPOULIDIS, Georgios. Polysemiotic communication vs. multimodality. *SAUC – Street Art & Urban Creativity Scientific Journal*, v. 5, n. 2, p. 26-31, 2019. DOI: [10.25765/sauc.v5i2.156](https://doi.org/10.25765/sauc.v5i2.156).

STEPHENS, Neil; RUIVENKAMP, Martin. Promise and ontological ambiguity in the in vitro meat imagescape: From laboratory myotubes to the cultured burger. *Science as Culture*, v. 25, n. 3, p. 327-355, 2016. DOI: [10.1080/09505431.2016.1171836](https://doi.org/10.1080/09505431.2016.1171836).

TEISER, Stephen F. *Reinventing the wheel: paintings of rebirth in medieval Buddhist temples*. Seattle, WA: University of Washington Press, 2007.

VELASQUEZ, Scott E. *Images of gender arrangements of three American social organizations: A content analysis of top ranked Google images of individuals performing organizational roles*. 2016. Ph.D. – Kansas State University, 2016.

WALSH, Wendy E. Investigating public perception of occupational therapy: an environmental scan of three media outlets. *American Journal of Occupational Therapy*, Bethesda, v. 72, n. 3, p. 1-10, 2018. DOI: doi.org/10.5014/ajot.2018.024513.

WASHBURN, Dorothy K.; CROWE, Donald W. (eds.). *Symmetries of culture: Theory and practice of plane pattern analysis*. Seattle, WA: University of Washington Press, 1988.

_____. (eds.). *Symmetry comes of age: the role of pattern in culture*. Seattle: University of Washington Press, 2004.

WEBB, Eugene J.; CAMPBELL, Donald T.; SCHWARTZ, Richard D.; SECHREST, Lee. *Unobtrusive measures*, rev. ed. Thousand Oaks, CA: SAGE, 2000.

WILDFEUER, Janina; PFLAEGING, Jana; BATEMAN, John; SEIZOV, Ognyan; TSENG, Chiao-I. *Multimodality, disciplinary thoughts and the challenge of diversity*. Berlin: de Gruyter, 2019. DOI: [10.1515/9783110608694](https://doi.org/10.1515/9783110608694).

WILKINSON, Mark D. *et al.* The FAIR guiding principles for scientific data management and stewardship. *Scientific Data*, v. 3, p. 1-9, 2016. DOI: [10.1038/sdata.2016.18](https://doi.org/10.1038/sdata.2016.18).

WISEMAN, Richard; OWEN, Adrian M. Turning the other lobe: directional biases in brain diagrams. *i-Perception*, v. 8, n. 3, p. 1-4, 2017. DOI: [10.1177/2041669517707769](https://doi.org/10.1177/2041669517707769).

WOLCOTT, Harry F. *Ethnography: a way of seeing*. 2nd ed. Lanham, MD: AltaMira, 2008.

WORRING, Marcel; SNOEK, Cees. Visual content analysis. In: LIU, Ling; ÖZSU, M. Tamer (org.). *Encyclopedia of database systems*. Boston, MA: Springer, 2009. p. 3360-3365. DOI: [10.1007/978-0-387-39940-9_1019](https://doi.org/10.1007/978-0-387-39940-9_1019).

WSPANIALY, Patrick; BROOKS, Justin; MOUSSA, Medhat. An image labeling tool and agricultural dataset for deep learning. *arXiv preprint arXiv:2004.03351*, 2020, 2020.

WU, Amy. Hillary's Heels: Examining gender and power through semiotics. In: PELKEY, Jamin; WALSH-MATTHEWS, Stephanie; SBROCCHI, Leonard (eds.) *Semiotics 2014: The Semiotics of Paradox*. Yearbook of the Semiotic Society of America. Ottawa: Legas, 2015. p. 473-490. DOI: [10.5840/cpsem201436](https://doi.org/10.5840/cpsem201436).

YU, Ning; WANG, Tianfang; HE, Yingliang. Spatial subsystem of moral metaphors: a cognitive semantic study. *Metaphor and Symbol*, v. 31, n. 4, p. 195-211, 2016. DOI: [10.1080/10926488.2016.1223470](https://doi.org/10.1080/10926488.2016.1223470).

ZLATEV, Jordan. Mimesis Theory, Learning, and polysemiotic communication. In: PETERS, Michael A. (org.). *Encyclopedia of Educational Philosophy and Theory*. Singapore: Springer, 2019. p. 1-6. DOI: [10.1007/978-981-287-532-7_672-1](https://doi.org/10.1007/978-981-287-532-7_672-1).

Comunicação secreta e história da criptologia: um desafio para as Humanidades Digitais

Peter Krapp¹

Tradução do inglês: Eduardo Harry Luersen

Resumo: Estudar a comunicação sob a perspectiva do sigilo soa paradoxal. Entretanto, pode-se afirmar seguramente que guardar é tão importante quanto compartilhar. Durante séculos a história da criptologia remeteu aos campos da linguagem, da tradução, da escrita e da interpretação, porém, desde a década de 1940 ela tem se baseado cada vez mais na matemática e na computação. No século XXI, a cultura digital depende da encriptação para que os sujeitos se comuniquem de maneira confiável e verificável com seus bancos, médicos, advogados e parceiros de negócios. Confrontar a questão da encriptação na atualidade demanda uma análise crítica das suas representações deturpadas. Indicamos que estamos trabalhando com Humanidades Digitais ao aplicarmos métodos das ciências de dados à comunicação estética. Então por que não o fazemos também quando aplicamos métodos das Ciências Humanas para pesquisar a cultura digital? Apesar de Galloway nos advertir que, na tarefa de processar e analisar informação quantitativa, “um pesquisador do campo da cultura que empregue tais métodos é pouco mais do que uma versão inferior da Amazon ou da Equifax”, é preciso que as Humanidades Digitais ajudem a reduzir a distância entre as abordagens quantitativas e qualitativas. Este artigo apresenta o projeto de um programa de ensino e a experiência pedagógica com uma turma de alunos de Ciências Humanas sobre a história da criptologia. Abordando a comunicação secreta desde tempos remotos até os dias de hoje, o artigo também explora como as mídias retratam as redes de computadores e como abordam o tema da cibersegurança.

Palavras-chave: Sigilo. Encriptação. Criptoanálise. Cibersegurança. Deturpações.

¹ Professor e coordenador do Departamento de Estudos de Cinema e Mídias na Universidade da Califórnia, em Irvine. Lecionou na Universidade de Minnesota e na Bard College e foi Pesquisador Visitante em Universidades na África do Sul, Alemanha, Brasil, Estados Unidos e Taiwan. É autor dos livros *Déjà Vu: Aberrations of Cultural Memory* (2004) e *Noise Channels: Glitch and Error in Digital Culture* (2011).

Secret communication and cryptologic history: a challenge for digital humanities

Abstract: It may seem paradoxical to study communication from the vantage point of secrecy, but arguably keeping is as important as sharing. For centuries, cryptologic history was about languages and translation, writing and interpretation, though since the 1940s it relies much more on math and computing. Digital culture in the 21st century relies on encryption for trustworthy and verifiable communications among people and their banks, doctors, lawyers, business partners. Addressing the stakes of encryption today requires a critical appreciation of its misrepresentations - digital humanities are invoked when we apply data-science methods to aesthetic communication, so why not when we apply Humanities methods to digital culture? Despite Galloway's warning that in parsing and processing quantitative information, "a cultural worker who deploys such methods is little more than a lesser Amazon or a lesser Equifax", digital humanities need to help bridge the gulf between quantitative and qualitative approaches. This paper lays out the course design for, and teaching experiences with, a class that introduces students in the humanities to the history of cryptology. Covering secret communication from ancient times to today, the paper also surveys how media portray computer networks and cybersecurity issues.

Keywords: Secrecy. Encryption. Cryptanalysis. Cybersecurity. Misrepresentation.

A história das mídias pode ser contada como a história das comunicações secretas – desde muito antes das primeiras transmissões de rádio e das suas interceptações, até muito depois da televisão, que uniu comercialmente a tecnologia militar ao entretenimento². Restam poucas dúvidas de que a cultura digital carrega, por todos os lados, as marcas dos debates sobre encriptação, segurança e confiança da comunicação entre os sujeitos e seus bancos, médicos, advogados e parceiros de negócios. Além disso, quando aqueles que se opõem a uma comunicação mais segura apelam a falácias sobre o “lado sombrio” da comunicação mediada por computadores, é preciso que enfrentemos o tema da encriptação partindo de uma análise crítica das suas representações deturpadas³.

As Humanidades Digitais não precisam implementar um uso estrito de ferramentas das chamadas ciências de dados para explorar as dimensões quantitativas dos seus objetos de interesse tradicionais; é perfeitamente possível utilizar uma abordagem interpretativa e manter a conversação das Humanidades Digitais com as Humanidades propriamente ditas. Na medida em que os mecanismos de busca prometem acesso em tempo real aos acervos das bibliotecas de pesquisa, em que as editoras acadêmicas se voltam à distribuição online e que instituições como a Biblioteca do Congresso Norte-Americano postam suas coleções na internet, os computadores passam a fazer parte do ensino e da curadoria da literatura, da história e da cultura audiovisual. A textualidade digital tem sido celebrada como uma nova forma de literatura, uma nova enciclopédia, uma biblioteca universal e como um metameio capaz de ingerir e substituir as mídias anteriores – mas também surgem novos formatos audiovisuais, que mobilizam nossa atenção a partir de modos de narração e interação distintos. A introdução da questão tecnológica nas Ciências Humanas altera o foco para as redes de tecnologias e instituições que per-

² Este artigo é uma versão atualizada e traduzida de uma publicação anterior: KRAPP, Peter. Beyond schlock on screen: teaching the history of cryptology through media representations of secret communications. *Proceedings of the 2nd International Conference on Historical Cryptology (HistoCrypt 2019)*, Linköping University Electronic Press, NEALT Proceedings Series 158:009, p. 79-85. Disponível em: <ep.liu.se/ecp/contents.asp?issue=158>. Acesso em: 25 mar. 2020.

³ Agradeço aos alunos dos meus seminários de escrita sobre a história das mídias da comunicação secreta da Universidade da Califórnia, em Irvine, das turmas de 2005, 2011, 2017 e 2019.

mitem que uma determinada cultura escolha, armazene e processe dados importantes. Também nos convida a explorar as frestas entre as práticas acadêmicas, como na suplementação dos modelos espaciais (texto, grafismos, ilustrações) com modelagens temporais destes dados (vídeos, modelos interativos), por exemplo. Ainda assim, há aqueles que entendem que a maioria das práticas das Humanidades Digitais não são propriamente Ciências Humanas. Para abrandar tais inquietações, as Humanidades Digitais devem construir formas de aplicar os métodos das Ciências Humanas à cultura digital, e não apenas utilizar as ferramentas digitais para exibir e representar temáticas tradicionais das Humanidades.

Podemos ilustrar esta controvérsia a partir de uma pesquisa sobre a comunicação secreta desde os tempos remotos até hoje. Esta costumava ser uma competência fundamental das Ciências Humanas: durante séculos, a história da criptologia se ocupou das linguagens e suas traduções e da escrita e sua interpretação. Evidentemente, desde a década de 1940 ela tem sido cada vez mais baseada na computação e na matemática – mas isto não significa que devemos abandonar completamente o tema aos engenheiros de software e teóricos dos números. Apesar de Galloway (2014, p. 110) nos advertir que, na tarefa de processar e analisar informação quantitativa, “um pesquisador do campo da cultura que empregue tais métodos é pouco mais do que uma versão inferior da Amazon ou da Equifax”, é preciso que as Humanidades Digitais ajudem a reduzir o abismo existente entre as abordagens quantitativas e qualitativas. Para testar a minha hipótese de que tanto a história como o atual debate sobre a segurança e a confiabilidade da informação seguem sendo preocupações fundamentais das Ciências Humanas, e que uma competência em Humanidades Digitais pode ser muito importante a esta dimensão, projetei e lecionei um curso que rastreia a longa história das comunicações secretas, e que simultaneamente extrapola, a partir de articulações históricas, diversas questões sobre a vida diante das condições técnico-midiáticas do século XXI.

Neste curso, os alunos desvendam uma história das mídias dos códigos e criptogramas de culturas remotas até o advento da computação, concentrando-se na comunicação secreta ainda anterior aos proto-computadores de Bletchey Park, e conectando os métodos históricos a questões contemporâneas sobre sigilo, privacidade e segurança na era da internet. A bibliografia inclui contos, artigos selecionados e capítulos de livros sobre a história da encriptação e da quebra de códigos. A cada semana também são realizados exercícios práticos (em sala de aula e lições de casa) e oficinas sobre aplicações contemporâneas de modelos históricos, com uma atenção particular aos apelos, na maioria das vezes essencialmente

inconciliáveis, por privacidade, segurança, confiança, integridade de dados e liberdade de expressão (DIFFIE e LANDAU, 2007). Todavia, por não se tratar de um curso de Ciências da Computação ou de Informática, não se demanda que os estudantes saibam trabalhar com aritmética ou que nutram um interesse pelos algoritmos. O arco geral do curso leva os alunos a considerarem as motivações de uma transmissão sigilosa e segura, que vão desde as garantias de integridade da comunicação às suas várias formas de autenticação. Isto permite contornar o equívoco ingênuo de se tomar ocultação e segurança como sinônimos. Uma vez tendo compreendido a diferença entre esteganografia e criptologia, as nuances entre as cifras de substituição e de transposição, os alunos estão aptos a realizar experiências práticas de decodificação a partir de exemplos históricos que vão desde a antiga cítala até a cifra maçônica. Após uma digressão sobre a esteganografia e a tinta invisível, é possível explorar os métodos de decodificação das cifras de substituição monoalfabética (KAHN, 1967; MACRAKIS, 2014; SINGH, 1999). A maior parte dos alunos da Universidade da Califórnia tem uma exposição anterior suficiente à história dos Estados Unidos, de modo que se torna interessante explorar as comunicações secretas da Guerra de Independência à Guerra Civil. Esta experiência é amplificada pela mostra de episódios selecionados de séries televisivas como *Turn* (2014-2017), sobre o *Culper Ring*⁴ – que também aparece no sexto episódio da quarta temporada (2012-2013) da série *White Collar: Crimes do Colarinho Branco*, ambientada na Nova Iorque contemporânea. Muitos alunos gostam de rastrear os processos de mecanização das cifras, desde o disco de Alberti até o disco do exército mexicano ou do disco de Jefferson ao cilindro M-94, utilizado pelo exército dos EUA até 1942.

As representações midiáticas do sigilo e da segurança da comunicação permitem que os alunos explorem a cifra de Vigenère, tentem observar o método Kasiki, visualizem as implementações de cifras desde o quadrado de Políbio até a cifra ADFGVX e compreendam o índice de coincidência de Friedman (KACKMAN, 2005; BAUER, 2013). Para introduzi-los à mística indelével das chamadas emissoras de números, eles escutam não apenas as gravações do Projeto CONET (1997), mas também ao álbum *Yankee Hotel Foxtrot* (2001) da banda Wilco, com a qual já têm certa familiaridade; além disso, o thriller *Códigos de Defesa* (2013), estrelado por John Cusack e Malin Akerman, conduz a uma produtiva discussão sobre o quão fiel às tecnologias reais o cinema e a televisão deveriam ser.

4 Nota do tradutor (N.T.): o *Culper Ring* foi uma organização de espões gerenciada pelo major Benjamin Tallmadge durante a ocupação britânica da cidade de Nova York, no apogeu da Guerra de Independência dos Estados Unidos (1775-1983).



Figura 1: Objetos de comunicação secreta apresentados na série *Turn*. Fonte: AMC. Disponível em: <amc.com/shows/turn/season-1>. Acesso em: 6 abr. 2020.



Figura 2: As emissoras de números representadas em *Códigos de Defesa*. Fonte: HFUnderground. Disponível em: <hfunderground.com/wiki/Numbers_stations_in_popular_culture>. Acesso em: 6 abr. 2020.

Em razão disso, um dos trabalhos de casa indispensáveis do curso é a escrita de uma resenha crítica sobre um audiovisual, enfatizando como as cifras e os códigos são retratados no cinema e na televisão. A lista de audiovisuais aceitáveis para este trabalho compreende obras que vão desde *Um Tenente Amoroso* (1935) e *Cipher Bureau* (1938) até *O Jogo da Imitação* (2014) e *Mr. Robot* (2015-2019). Porém, a maioria dos alunos não costuma escolher os exemplos mais antigos nem os mais recentes, optando seguidamente por filmes de espionagem dos anos 1980 ou 1990, como *Quebra de Sigilo* (1992) e *Pi* (1998).

O curso trabalha com narrativas e suas visualidades e, como não requer conhecimentos avançados de matemática nem de informática, explora principalmente aquelas cifras e códigos anteriores ao advento da computação, sem excluir os códigos *Navajo*, a máquina *Enigma*, o computador *Colossus*, dentre outros. Os alunos logo entendem porque os linguistas desenvolveram a análise de frequências como uma forma para decodificar cifras de substituição simples, e como os nomencladores e os livros de códigos auxiliaram no comércio e na diplomacia (KAHN, 1967). Por não exigir que os alunos do curso saibam realizar fatoração primária ou funções de dispersão criptográfica (funções *hash*), é de fundamental importância fomentar discussões sobre como o seu uso da internet é dependente de princípios criptográficos – uma vez que eles tenham partilhado o quanto as suas vidas cotidianas estão em torno da confiança nas comunicações online com bancos e lojas, médicos e farmácias, instituições educacionais e formas de entretenimento, fica fácil ilustrar o quanto a infraestrutura da segurança das comunicações depende de criptogramas assimétricos (BAUER, 2007; QUISQUATER, 1990).

O curso é concluído com um exame de quebra-cabeças criptológicos sem resolução, abordando o Manuscrito Voynich, as cifras de Beale, o enigma do Zodíaco e o *Kryptos* (CLEMENS, 2016; BAUER, 2017; SCHMEH, 2015). Ao terminarem este seminário de escrita intensiva, espera-se que os alunos tenham desenvolvido: um vocabulário crítico e histórico relevante; um conhecimento mais aprofundado sobre a história das mídias; uma apreciação mais perspicaz dos códigos e criptogramas; a habilidade para analisar criticamente as reivindicações conflitantes sobre a comunicação; e um melhor entendimento do sigilo ficcional e real (GLASS, 2013; KOBLITZ, 2010; KOSS, 2014). Além de familiarizar os alunos a conteúdos conceituais e históricos, este curso envolve o desenvolvimento de habilidades avançadas de literacidade da informação, a partir da procura, avaliação e integração de informações de fontes diversificadas para a escrita específica exigida pela disciplina. Ainda assim, a principal dificuldade de um curso como este se trata de como superar o paradigma do *schlock* – o fato irrefutável de que a maioria das representações midiáticas das comunicações secretas são batidas, estereotípicas, deturpadas, enganosas ou simplesmente erradas. É sob este aspecto que a exposição ao uso prático e experimental das ferramentas digitais se torna uma necessidade pedagógica.

Qual é a aparência da comunicação encriptada? O problema com as representações audiovisuais da cibersegurança, em particular, e das redes de computadores, em geral, é que com grande frequência elas são transformadas, na tela, em caricaturas extravagantes. Mesmo as séries televi-

sivas em que a computação é central, como *CSI Cyber* (2015), interpretam inúmeros detalhes de forma tão equivocada que são poucos os telespectadores com alguma experiência em informática que aguentam assisti-las. A computação não se trata de pixels e luzes piscando – e em nada ajuda inundar o roteiro com um jargão mal utilizado e pronunciado. Os códigos nocivos não piscam em vermelho na sua tela, e uma análise forense da máquina demora mais do que alguns minutos. Em sua maior proporção, os cibercrimes se ocupam do “*phishing*” de números de cartões de crédito e de segredos comerciais, e apenas muito raramente de sequestros. Ao tratarem dos riscos online reais e imaginários, os filmes e as séries de tv costumam entregar repetições desgastadas, que são nocivas não apenas por perpetuarem os estereótipos do *hacking* como um flerte adolescente (e normalmente masculino) com o crime, como também por representarem a forma do “espaço” de dados como se fosse um jogo de fliperama.

Mesmo que este seja um problema global, meus exemplos audiovisuais têm um viés anglófono e centram-se nos EUA; mas poderíamos substituí-los por exemplos vindos de muitos outros lugares. Os incontáveis filmes e séries norte-americanos que se equivocam quanto à computação, encriptação e deciptação, tendem a cometer sobretudo dois tipos de erros: a glamorização das ações de uma pessoa em frente a um computador e a tentativa de tornar visível o fluxo de dados em rede, normalmente de maneiras extravagantes. Tomemos como um exemplo inicial o filme *Hacker* (2015), de Michael Mann. Nele, a personagem de Chris Hemsworth sai da cadeia para ajudar no combate internacional ao cibercrime. Ignoremos a ideia de que o Thor deveria ser um nerd em computação (e o fato de que em sua cela ele lê Baudrillard e Derrida, ao invés de, digamos, Schneier ou Kahn). Na trama em que um empregado do governo deve ser ludibriado a trocar a sua senha (para que um intruso possa ganhar o acesso), ao tentar dar visibilidade aos dados em rede, a obra regressa diretamente à terrível época de filmes como *A Rede* (1995), *Hackers: Piratas de Computador* (1995), *Quebra de Sigilo* (1992), *Jogos de Guerra* (1983) ou *Tron: Uma Odisseia Eletrônica* (1982). É verdade que *Hackers: Piratas de Computador* é mais lembrado por apresentar Angelina Jolie como uma das duas estudantes que se envolvem em extorsão corporativa. E também é verdade que os protagonistas passam mais tempo falando bobagens e fazendo festa do que usando computadores, mas além disso este filme exhibe um vírus que fala e que tem um rosto. Já *A Rede* mostra Sandra Bullock navegando freneticamente por entre sistemas de *Bulletin Board*, como se pedir uma pizza online fosse um ato radical e subversivo. Ainda

que mostre alguns aspectos verdadeiros da rede (como os endereços de IP), o filme não faz o suficiente para fundamentá-los na realidade computacional – não era possível conectar-se a um e-mail através da *Telnet*, por exemplo. Seria necessário um endereço de IP e um número de porta TCP: o *login* de um e-mail só poderia ser iniciado após estar conectado. Também não seria possível que um vírus de um *Macintosh* de 1995 infectasse um computador mainframe.

Curiosamente, *Quebra de Sigilo* (1992), uma ficção criminal concebida durante a produção de *Jogos de Guerra* (1983), apresenta uma caixa preta capaz de quebrar qualquer tipo de criptografia computacional, ameaçando desestabilizar a economia mundial (o que já suscitava o tema da criptografia pós-quântica). Ambos os filmes tendem a se precipitar ao lidarem com a tecnologia computacional – ainda que seja possível alterar as notas de alguém no colégio a partir de um computador doméstico caso a escola seja descuidada, é impossível lançar um míssil balístico intercontinental (MBI) a partir da mesma máquina, usando apenas o mesmo modem de acesso discado. *Tron: uma Odisséia no Espaço* (1982) é marcante pelo uso pioneiro que faz da computação gráfica, mas a ideia de que se poderia entrar na rede e operá-la como se fosse a malha gráfica de um videogame teve uma influência perniciosa sobre o cinema e a televisão.

Em *A senha: Swordfish* (2001), um filme extravagante com John Travolta, encontramos uma das mais engraçadas transgressões do cinema contra a computação. No filme, o personagem de Hugh Jackman é forçado a acessar remotamente um computador do Departamento de Defesa dos Estados Unidos (supostamente através de um antigo computador mainframe *PDP10* localizado em um porão do Instituto de Tecnologia da Califórnia). Ele consegue fazê-lo batendo palmas espalhafatosamente e digitando muito rápido, enquanto é ameaçado por capangas e sexualmente coagido por uma jovem... E obviamente ele obtém o acesso em 85 segundos, mesmo tendo recém superado uma condenação que o obrigou a ficar por um longo período sem usar computadores; e é óbvio que o seu software antigo, que ficou armazenado numa fita em algum porão, ainda funciona como um *malware* destrutivo, ainda que graficamente ele pareça muito ultrapassado... Em *Superman 3* (1983), pelo menos o personagem de Richard Pryor podia se concentrar para hackear um satélite meteorológico, por mais improvável que o fosse fazê-lo utilizando uma linguagem *BASIC*, com os comandos *PRINT* e *LIST* – sem contar outras façanhas realizadas por Pryor na sua função de assistente de computação, como alterar dados de folhas salariais e desordenar semáforos.



Figura 3: Richard Pryor em *Superman 3*. Fonte: *Internet Movie Database*. Disponível em: <imdb.com/title/tt0086393>. Acesso em: 6 abr. 2020.



Figura 4: Cena de *hacking* em *A Senha: Swordfish*. Fonte: Packet Pushers. Disponível em: <packetpushers.net/a-fun-look-at-workstations-interfaces-for-it-folks>. Acesso em: 6 abr. 2020.

Infelizmente, as coisas não ficaram muito mais sofisticadas com o passar dos anos – pensemos, por exemplo, na série contemporânea *Homeland* (2011-2020): quem poderia imaginar que o servidor da CIA que os dois net-ativistas de Berlim encontram por acaso (na quinta temporada, de 2015) lhes daria acesso a um diretório repleto de arquivos, cujos longos nomes *todos* contêm a cadeia de caracteres “CIA”? Todas as listas de arquivos que você utiliza levam o nome do seu empregador? Deveríamos supor que arrancar um cabo físico da parede é a única coisa que os

peritos de Langley poderiam fazer para defender a CIA de uma enxurrada massiva de pesquisas por *cam-shows* pornô na internet? Programas como estes tendem a tratar a computação com menos seriedade do que Indiana Jones trata a arqueologia.

Isto não se trata apenas de uma questão de verossimilhança e realismo. Ainda que sabidamente o autor de ficção científica Arthur Clarke tenha declarado que qualquer tecnologia suficientemente avançada se torna indistinguível da magia, esta situação apoia-se em uma ignorância generalizada, que é legada de eras pré-letradas e é incompatível com os objetivos da educação. A magia pode ser aceita em uma ficção científica, mas não em uma universidade; nosso interesse se concentra em conceitos aplicáveis. Portanto, por exemplo, um aluno deste curso que queira discutir Harry Potter deve levar em consideração um sistema de identificação por dois fatores – para ter acesso às salas comunais de Hogwarts é necessária uma combinação de algo que você tenha (por exemplo, uma varinha) ou que você seja (não ser um “trouxa”) com algo que você saiba (uma senha); sem mencionar a ofidioglossia (língua das cobras) que é necessária para acessar a câmara secreta, ou o sacrifício de sangue exigido para adentrar a caverna onde se esconde uma *Horcrux*... Como os filmes retratam o acesso restrito ao Caldeirão Furado e à plataforma 9 $\frac{3}{4}$ da estação *King's Cross*? Quem tem acesso (e como) ao banheiro dos monitores-chefes e ao escritório do professor Dumbledore? Embora estas perguntas possam soar extravagantes, são questões como estas que estão em jogo.

Antes da série *Mr. Robot* (2015-2019), raramente a televisão retratou questões de segurança da informação de maneira realista. *Mr. Robot* é sobre atividades de uma empresa de cibersegurança. Os códigos que a série mostra nas telas dos computadores são reais e não há efeitos sonoros cafonas ou arroubos de fantasia. Ao representar o hack de um aparelho celular com *Android*, pelo uso de um chip que executa um gerenciador de inicialização, a série se refere a tecnologias reais – neste caso, especificamente, ao software *Flexispy*. Notavelmente, *Mr. Robot* não se abstém de retratar os roteadores *TOR*, um ataque distribuído de negação de serviço (*DDoS*) em servidores corporativos e a instalação de *malwares* – neste caso, um *trojan* de acesso remoto que lembra um software verdadeiro chamado *DarkComet*.

Ironicamente, em meio à sua trama conspiratória, o pastiche sci-fi *Matrix Reloaded* (2003) é um dos poucos filmes em estilo *schlock* a mostrar uma cena realística: por uma vez ao menos evita a mania habitual de dar visualidade ao ciberespaço como se fosse um voo vertiginoso através dos desfiladeiros sombrios de uma Manhattan de dados mal renderiza-

dos. O filme mostra Trinity (nem masculina nem adolescente, mas interpretada por Carrie-Anne Moss) trabalhando com o teclado, em vez de representá-la operando alguma engenhoca qualquer com uma interface futurista. E, para encontrar vulnerabilidades na rede elétrica, ela utiliza um tipo de software existente: o *NMAP*, um scanner de portas *TCP* e *UDP* conhecido por gestores de sistemas no mundo todo, que é executado através de linha de comando.



Figura 5: A personagem Trinity em *Matrix: Reloaded*. Fonte: *The Killtime*. Disponível em: <thekilltime.com/the-matrix-1999>. Acesso em: 6 abr. 2020.

O ciberthriller alemão *Invasores: Nenhum Sistema está Salvo* (2014) apresenta de forma semelhante o uso destes softwares (*exploits*) na rede elétrica – remotamente, o protagonista Benjamin procura afetar um serviço local utilizando um *script* que parece dotado de poderes universais quando digitado em uma interface de linha de comandos. No thriller *Ultimato Bourne* (2007), a CIA hackeia o servidor de e-mail de um jornal britânico e a tela mostra um uso realista de *SSH*, de servidores *Postfix* com protocolos *SMTP* e de um servidor de nomes de domínio (*DNS*) em uma interface do *Unix*. No mesmo ano, outra franquia de *thrillers* apresentava um *exploit* interessante logo nos primeiros dez minutos de filme: para combater um ciberterrorista, o protagonista de *Duro de Matar 4.0* (2007) se junta a um jovem hacker. Também nos seus primeiros dez minutos, o filme sueco *Os Homens que Não Amavam as Mulheres* (2009), adaptação para as telas do livro de Stieg Larsson, mostra as habilidades computacionais de sua protagonista. Algo que o remake hollywoodiano, de 2011, não faz. No filme de super-heróis *Quarteto Fantástico* (2015, baseado nos qua-

drinhos da *Marvel*) ao menos é possível ver Sue Storm (personagem de Kate Mara) rastreando um companheiro online: sua tela pisca “*IPSCAN*”, “*TRACEROUTE*” e “*PORTSCAN*” – é muito raro ver representações realistas de tecnologias de rede.

Evidentemente, enquadrar corretamente as tecnologias computacionais no audiovisual não é apenas uma questão de software e hardware; a cibersegurança também envolve engenharia social – a exploração de padrões de comportamento, oportunidades e vulnerabilidades. Ainda que os fabricantes de hardware se aproximem dos produtores de cinema e televisão para exibir os seus artefatos, tanto a indústria de software quanto o setor educacional não cansam de perder oportunidades de mostrar a computação como algo interessante, estimulante e desafiador – sem dissimulá-la. De fato, na cultura popular o *hacking* não é mais celebrado como a atividade normalmente inócua (e ocasionalmente muito lucrativa) de alguns entusiastas da computação. A televisão parou de romantizar as obsessões de nerds engenhosos e a imprensa já não fala mais tanto do “espírito virtuoso” do capitalismo digital. Ao invés disso, os jornalistas se ocupam de vender os aspectos sinistros do *hacking* como uma ameaça sistêmica e irredutível das mídias digitais. Pouco importa que até o final dos anos 1980 um hacker era apenas alguém que, sem a ajuda de manuais e através de tentativa e erro, acabava adquirindo a habilidade de operar bem os computadores. Foi só alguns anos depois que os comentaristas começaram a temer que o *hacking*, quando usado maliciosamente, poderia se configurar como um problema sério e oneroso. Na maior parte do tempo, a cultura digital foi centrada no acesso, no aprendizado, na privacidade e na liberdade de expressão (BAMFORD, 1982; LEVY, 2001; SCHNEIER, 2004). Ainda assim, por uma mudança paradigmática na opinião pública, assim como nas políticas econômicas e em ações legais relacionadas ao ensino e às tecnologias em rede, os comentaristas mais alarmistas passaram a demonizar qualquer um que tentasse acessar mais do que aquilo que a interface oficial e limitada tinha a oferecer. Tudo isso conduz ao questionamento sobre como poderiam ser elaboradas tarefas apropriadas para os alunos do curso.

Cultuar o sigilo poderia facilmente nos levar a um ressurgimento global de rumores irracionais e, infelizmente, isto é o que se vê efetivamente em boa parte da cultura da internet. O nosso futuro tende a ser consideravelmente empobrecido quando as teorias da conspiração assumem o lugar de uma cultura computacional crítica. Seguramente, ensinar os conceitos básicos de cibersegurança e algumas linhas gerais da

história da criptologia aos alunos de cinema e de estudos de mídia pode aumentar, ainda em tempo, as chances de que os seus roteiros e cenas produzam representações audiovisuais mais precisas e inteligentes da computação e da segurança da comunicação.

Para aguçar a percepção dos alunos tanto em relação às representações problemáticas da cibersegurança como quanto às possibilidades plausíveis e persuasivas de visualização da criptografia, requer-se que eles escrevam sinopses argumentativas para filmes e pilotos de tv, a partir de contos que são distribuídos entre eles. Ainda que no jargão da indústria uma sinopse seja diferente de um argumento completo (uma sinopse condensa brevemente o tom da narrativa, enquanto o argumento vai até os aspectos essenciais da representação audiovisual da estória), o que importa para os propósitos pedagógicos do curso é que este documento destaque os elementos necessários com algum ritmo e estilo cinematográfico e passe uma primeira impressão sobre as personagens, a aclimação e os recursos visuais utilizados para evocar um tempo e um espaço. Esta sinopse argumentativa não deve apenas recontar a estória; ela deve ser suficiente para permitir que sejam tomadas decisões a partir da avaliação tanto das ideias como da solução audiovisual descrita. Além de demarcar os *beats* da narrativa com uma atenção especial à representação audiovisual dos aspectos da segurança da comunicação, o trabalho deve incluir um título e uma premissa, introduzir as personagens principais, definir os cenários, dramatizar os principais conflitos que conduzem ao clímax e esboçar uma resolução dramática. Diferente de um conto, a sinopse argumentativa não deve dizer, mas mostrar o que um personagem está pensando; não deve fornecer todo o contexto, mas esboçar alguns diálogos; por motivos pedagógicos (e porque, de fato, escrever diálogos é difícil), o objetivo neste ponto é descrever mais sucintamente alguns diálogos que o roteiro pleno iria apresentar. Os contos distribuídos entre os alunos são seleções de histórias de detetive americanas e britânicas que datam da virada do século XIX para o século XX. Ao transformarem estes materiais obsoletos em uma sinopse argumentativa, os alunos não apenas recontam uma peça de ficção. Estando particularmente atentos às formas como a comunicação secreta pode ser apresentada na tela, eles acabam atualizando e remodelando os *beats* destas narrativas aos seus próprios gostos contemporâneos.

Além disso, os alunos devem compilar materiais de referência ao modo de verbetes de enciclopédia, referindo-se a determinados nomes e conceitos importantes à história da criptologia. Esta tarefa consiste em realizar uma pesquisa (com pelo menos quatro ou cinco referências) para

definir, descrever e discutir os termos elencados ou as pessoas escolhidas, explicando claramente como cada verbete se relaciona com os tópicos do curso. Pelo menos uma das fontes de referência deve ser proveniente de um repositório online para pesquisas acadêmicas (tais como o *JSTOR* ou o *Project MUSE*), para que os alunos possam ir aos poucos se familiarizando aos sistemas de biblioteca que fazem parte do trabalho acadêmico. Podendo se ocupar de nomes que vão desde Alberti e Trithemius até Diffie e Schneier, e de conceitos que incluem o princípio de Kerckhoffs, as provas de conhecimento-zero, o *Atbash*, o *PGP* e o chip *Clipper*, eventualmente os alunos encontram dificuldades para sintetizar suas descobertas em um único verbete que seja conciso, mas ainda suficientemente abrangente. Uma avaliação em sala de aula é realizada na metade do período letivo, contendo tanto questões de múltipla escolha sobre fatos históricos como algumas questões abertas que requerem alguns parágrafos de reflexão. De que modo as palavras-chave podem ser utilizadas para aprimorar uma cifra de César? Qual é o segundo trígama mais comum na língua inglesa? Quais tipos de tintas invisíveis você saberia citar? Qual é o nome do antigo método grego para garantir a segurança das mensagens confidenciais? Qual diplomata dos Estados Unidos está ligado aos discos cifrados? Caso não se tenha a senha, qual é o primeiro passo para começar a decodificar uma mensagem cifrada com o método de Vigenère? O que ocasionou a reviravolta estatística de Babbage na criptoanálise, e porque não foi ele, mas Kasiski quem a publicou? Como parte da avaliação, consta ainda um artigo substancial a ser escrito em casa. A escolha individual do tema e os primeiros esboços devem ser trabalhados em aula. Ao menos uma das fontes deve ser proveniente de um banco de dados online de pesquisa acadêmica, novamente, e o tema do estudo deve ser relacionado aos materiais do curso. Os governos deveriam ter a permissão para acessar a comunicação encriptada de qualquer indivíduo para prevenir crimes ou as empresas de tecnologia deveriam empregar a encriptação mais segura possível para proteger a privacidade dos usuários? Quais são as formas contemporâneas de esteganografia e quão práticas elas são na proteção e/ou transmissão de informações sigilosas através de arquivos superficialmente inalterados de áudio, de imagem, de vídeo etc.? Em todos os processos de escrita deste curso, tanto em aula como em casa, os alunos realizam avaliações por pares de diferentes esboços, através de um ambiente online compartilhado. Desta forma, ao serem avaliados, os trabalhos finais já não contêm mais as falhas e imprecisões comuns aos primeiros esboços.

Os alunos logo descobrem por conta própria o motivo pelo qual as representações audiovisuais das comunicações secretas são tão frequentemente errôneas e afetadas. Mas são poucos aqueles que conseguem encontrar soluções suficientemente criativas a ponto de serem exaltadas por seus colegas nas revisões por pares. Muitas das atividades deste curso não demandam resenhas ou comentários sobre a bibliografia, diferentemente das abordagens críticas e analíticas que geralmente são exigidas dos estudantes das Ciências Humanas. Quando a escrita objetiva contar uma história por meios audiovisuais, não basta relatar os pensamentos de uma personagem nem descrever os contextos técnicos e biográficos. Decerto, não se pode esperar uma escrita verdadeiramente criativa – os alunos sabem que não devem copiar nem inventar completamente os diálogos e as personagens de suas sinopses argumentativas. Uma vez que eles percebem como é possível manter a fidelidade à dimensão conceitual de um conto e ainda assim contribuir com ideias consideravelmente inventivas para um roteiro baseado na mesma obra, os alunos tendem a entregar-se com uma engenhosidade impressionante.

Por outro lado, mesmo que eles apreciem descobrir como escrever seus próprios nomes usando uma cifra maçônica ou que consigam expressar as diferenças entre as cifras de substituição e transposição, não se deve esperar que estudantes de Ciências Humanas instalem o *JCrypt* em seus computadores, ou que estudem a fundo a matemática envolvida na criptografia assimétrica (KOBBLITZ, 1997; KAUR, 2008; WINKEL, 2008; KURT, 2010). Entretanto, certamente se pode confiar no seu envolvimento com uma ementa voltada a um levantamento da história da comunicação secreta que, mesmo tratando predominantemente de estágios anteriores à computação, permite que eles tirem conclusões sobre suas próprias vidas no século XXI. Mesmo (e especialmente) aqueles alunos que não forem cientistas da computação promissores devem ser capazes de debater o papel da criptografia durante a Segunda Guerra Mundial, tendo lido sobre as máquinas *Enigma* e *Purple*, sobre Alan Turing e os matemáticos poloneses que fizeram descobertas seminais naquele período. Os estudantes invariavelmente demonstram interesse na cibersegurança contemporânea, tanto em aspectos do gerenciamento de senhas quanto em questões sobre a confiança depositada nas diversas plataformas de comunicação, sejam elas *WhatsApp*, *Messenger*, *Telegram* ou *Signal*. Eles tendem a ser entusiastas das redes sociais e, em geral, muito menos céticos

em relação à mineração de dados e à publicidade digital do que sugerem as sondagens nacionais (PEW RESEARCH CENTER, 2017a). Eles se animam com o Telegrama Zimmermann, mas muito mais tentando imaginar o que seriam os *casus belli* equivalentes na atualidade. Eles tendem a ser bastante céticos em relação a algumas das afirmações de David Kahn sobre o porquê de os alemães terem perdido a Segunda Guerra Mundial para a inteligência norte-americana; não vilanizam nem veneram figuras como Julian Assange ou Edward Snowden, e suas discussões e trabalhos de casa mostram uma clara preferência por debater a segurança dos servidores de jogos digitais em vez da segurança das operadoras de cartão de crédito, ou a confiança nas plataformas de mídias sociais em vez das instituições políticas.

Os discentes tendem a ter discussões acaloradas e bem informadas sobre questões de identidade, anonimato e pseudônimos online. Descontentes em ficar apenas com o ponto de vista da maioria, alguns dos alunos mantêm durante todo o curso fortes reservas quanto a certos conflitos entre privacidade e segurança pública, confiança e tecnologia avançada, liberdade de expressão online e responsabilidade pessoal. Em suas sinopses argumentativas, ao atualizarem a ficção policial clássica, como o conto de Sherlock Holmes sobre os dançarinos ou um conto de Isaac Asimov sobre encriptação, os alunos demonstram criatividade, imaginando na forma do graffiti um código que permanece secreto mesmo estando visível a todos. Talvez isto não surpreenda, dado que eles possuem menos livros impressos e leem mais arquivos digitais em seus diversos dispositivos. Menos alunos se interessam por códigos escondidos em livros, tal como os que constam na segunda temporada da série *Burn Notice* (2008-2009) (uma bíblia roubada ocupa todo o enredo desta temporada), no filme *A Lenda do Tesouro Perdido* (2004) (onde as coordenadas no verso da Declaração da Independência levam ao suposto tesouro) ou no mistério *O Vale do Terror* (1915), de Sherlock Holmes, apesar da longa popularidade do personagem na televisão e no cinema. Também não surpreende que os estudantes estejam menos interessados em códigos que envolvam mais de um idioma, mesmo que apreciem a interrelação de longa data da história da criptologia com áreas como a Tradução e a Filologia e não apenas as suas formas mais populares, como a pedra de roseta ou os processos de descodificação das mensagens alemãs e japonesas na Segunda Guerra Mundial.

Da mesma forma, os alunos tendem a mostrar uma revigorante falta de respeito pelos aspectos antiquados dos contos que se encarregam de atualizar: em vez das caixas de rapé, canetas e lenços envenenados, as suas versões destas narrativas trazem *smartphones*, boates e graffiti; em

vez de retratarem personagens escandalizadas por casos de infidelidade ou fraudes fiscais, os enredos imaginados incluem gafes em redes sociais e concursos de camisetas molhadas. Os ensaios finais entregues pelos alunos demonstram que as metas do curso foram alcançadas, pela constatação de um evidente domínio das dimensões históricas e conceituais da criptologia. É comum o recebimento de vários artigos abordando questões de falsidade ideológica e como se prevenir contra ela. Normalmente são apresentados argumentos pertinentes tanto a favor como contra o acesso do governo a *backdoors* de criptografia; alguns estudantes mais ambiciosos e com conhecimentos de informática também se dispõem a quebrar a cabeça com criptogramas ainda não resolvidos (SCHMEH, 2015; BAUER 2017) ou a enfrentar o desafio de demonstrar como os princípios de Auguste Kerckhoffs (de 1883) ainda são válidos na cultura das mídias móveis de hoje.

Para introduzi-los a uma diversidade de implementações técnicas das questões estudadas no curso, os alunos também realizam breves exercícios semanais, que vão desde a instalação correta de *browsers* e redes virtuais privadas até a comparação entre gerenciadores de senhas, passando ainda pela autenticação multifator e pelos meios de fazer uma utilização segura das redes sociais. Agora que a universidade em que eles estão matriculados passou a utilizar sistemas que requerem autenticação multifator, eles estão se dando conta de que apenas o uso de senha não é suficiente para impedir o acesso não autorizado a dados individuais e institucionais; e a maioria dos alunos tende a levar a proteção de suas informações de identificação pessoal bastante a sério. Atualmente, mesmo no ambiente acadêmico as senhas são frequentemente expostas pelo uso de malwares, pelos ataques criptográficos de força bruta, *phishing* e outros *exploits*, o que torna a história e o futuro da segurança da comunicação uma fonte de novas perguntas importantes para a vida cotidiana dos alunos. Em parte, o que este curso propõe é que os estudantes descubram estes princípios básicos por conta própria, sem que o seu comportamento seja induzido, como normalmente as instituições tendem a fazer.

A integração de recursos online ao curso também ajuda a contextualizar esta situação. Uma conta no *Instagram* é um repositório de imagens da história da criptografia mais prontamente acessível do que um projetor de slides costumava ser. Além de poderem verificar diretamente pelo calendário da Fundação Museu Nacional da Criptologia⁵ o que aconteceu em determinada data na história da criptologia, os alunos também podem utilizar conversores e leitores de código

5 N.T.: *National Cryptologic Museum Foundation* (NCMF). Ver: <cryptologicfoundation.org>. Acesso em: 5 abr. 2020.

morse, brincar com tipografias *Pigpen* e com geradores de anagramas. Também podem usufruir do aparato de treinamento em cibersegurança da universidade de modo muito mais informado. Me reuni por diversas vezes com os especialistas em TI do campus, responsáveis por divulgar formas seguras de agir na internet. Quer seja nas apresentações expositivas seguidas de discussão ou em mesas-redondas de regime mais concentrado, é perceptível que aqueles alunos que fizeram o curso foram melhor avaliados nos módulos de treinamento em cibersegurança da universidade do que o público geral, de acordo com enquetes relevantes (PEW RESEARCH CENTER, 2017b).

Referências

- BAMFORD, James. *The puzzle palace: inside the National Security Agency*. New York: Penguin, 1982.
- BAUER, Craig P. *Secret history: the story of cryptology*. Boca Raton: CRC Taylor & Francis, 2013.
- BAUER, Craig P. *Unsolved! The history and mystery of the world's greatest ciphers from Ancient Egypt to online secret societies*. Princeton: Princeton University Press, 2017.
- BAUER, Friedrich L. *Decrypted secrets: methods and maxims of cryptology*. Berlin: Springer, 2007.
- CLEMENS, Raymond. *The Voynich Manuscript*. Yale: Yale University Press, 2016.
- DIFFIE, Whitfield; LANDAU, Susan. *Privacy on the line: the politics of wiretapping and encryption*. Cambridge: MIT Press, 2007.
- GALLOWAY, Alexander. The cybernetic hypothesis. *Differences: a journal of feminist cultural studies*, n. 25, v.1, 2014, p. 107-130.
- GLASS, Darren. A first-year seminar on cryptography. *Cryptologia*, n. 37, v. 4, 2013, p. 305-310.
- KAHN, David. *The codebreakers: the story of secret writing*. New York: Macmillan, 1967.
- KACKMAN, Michael. *Citizen spy: television, espionage, and Cold War culture*. Minneapolis: University of Minnesota Press, 2005.
- KAUR, Manmohan. Cryptography as a pedagogical tool, *PRIMUS*, v. 18, n. 2, 2008, p. 198-206.

KOBLITZ, Neal. Cryptography as a teaching tool. *Cryptologia*, n. 21, 1997, p. 317-326.

KOBLITZ, Neal. Secret codes and online security: a seminar for entering students. *Cryptologia*, n. 34, 2010, p. 145-154.

KOSS, Lorelei. Writing and information literacy in a cryptology first-year seminar. *Cryptologia*, n. 38, 2014, p. 223-231.

KURT, Yesem. Deciphering an undergraduate cryptology course. *Cryptologia*, n. 34, v. 2, 2010, p. 155-162.

LEVY, Stephen. *Crypto*. London: Penguin Books, 2001.

MACRAKIS, Kirstie. *Prisoners, lovers, and spies: the story of invisible ink from Herodotus to al Qaeda*. Yale: Yale University Press, 2014.

SCHMEH, 2015. *Klaus Schmeh's List of Encrypted Books – Klausis Krypto Kolumne*. Disponível em: <scienceblogs.de/klausis-krypto-kolumne/klaus-schmehs-list-of-encrypted-books>. Acesso em: 6 abr. 2020.

SCHNEIER, Bruce. *Secrets and lies: digital security in a networked world*. Indianapolis: Wiley, 2004.

SINGH, Simon. *The code book: the science of secrecy from Ancient Egypt to quantum cryptography*. New York: Anchor Books, 1999.

PEW RESEARCH CENTER. *Younger Americans express more support for encryption than older adults*. 2017a. Disponível em: <pewresearch.org/internet/2017/01/26/3-attitudes-about-cybersecurity-policy/pi_01-26-cyber-03-00>. Acesso em: 6 abr. 2020

PEW RESEARCH CENTER. *Cybersecurity Knowledge Quiz*. 2017b. Disponível em: <pewresearch.org/internet/quiz/cybersecurity-knowledge>. Acesso em: 6 abr. 2020

QUISQUATER, Jean-Jacques *et al.* How to explain zero-knowledge protocols to your children. In: BRASSARD, Gilles. (ed.) *Advances in cryptology. Crypto 89*, LNCS 435, 1990, p. 628-631.

WINKEL, Brian. Lessons learned from a mathematical cryptology course. *Cryptologia*, n. 32, v. 1, 2008, p. 45-55.



EXTRA DOSSIÊ

A dialogue on cognitive semiotics: minds, and machines¹

Vincent Colapietro²

Winfried Nöth³

Edited by Guilherme Henrique de Oliveira Cestari⁴

dx.doi.org/
10.23925/1984-3585.2020i21p167-184

W.N.: *Technologies of Intelligence and Digital Design*, abbreviated TIDD, the interdisciplinary postgraduate program of the Catholic University of São Paulo, founded by Lucia Santaella, was inaugurated in 2006. In its 15th year of existence, it is offering a number of series under the name “TIDD digital”, each dealing with topics from the program’s various study fields. We have had series on artificial intelligence and art, artificial intelligence and law, artificial intelligence and ethics, and on intelligent business technologies. Today, I am pleased to introduce a new series on *cognitive semiotics*. It is divided into two episodes. Today, in English, the question is “Cognitive semiotics: What is it?”

The interdisciplinary MA and PhD program TIDD deals with the contemporary challenges posed by the digital technologies. It counts on the cooperation of the following interdisciplines: computation, information technologies, social network, cognition and education, design and esthetic technologies, and artificial intelligence. Today’s focus is on cognition, a research field of an interdiscipline of its own, called cognitive science. Collaborators of this research field are scholars from the disciplines of cognitive psychology and philosophy, linguistics and anthropology, computer science, and artificial intelligence.

¹ The dialogue took place online on the channel @TIDDigital (linktr.ee/TIDDigital), on July 24, 2020, as Episode 1 of the Series Semiótica Cognitiva under the title “Cognitive Semiotics – What is it?” <facebook.com/events/3287017414683647/>.

² University of Rhode Island, Department of Philosophy, Kingston, RI, USA.

³ PUC-SP, TIDD, São Paulo, Brazil.

⁴ PhD – TIDD (PUC-SP). CV Lattes: lattes.cnpq.br/7987513983124155.

Why should cognition be a topic of research for students whose professional background and perspectives are in areas such as computation, digital design, social networks, robots, or artificial intelligence? Cognition has to do with perception (seeing, hearing, feeling, etc.), thinking, reasoning, memory, learning, emotion, and consciousness. TIDD cannot have the ambition to address all of these issues. Instead, its focus is on questions concerning cognition in more specific respects. How does our daily interaction with intelligent technologies influence or enhance our capacities of cognition? Do computers, robots and other so-called technologies of intelligence perceive? Do they think? May they have beliefs? Are these technologies intelligent at all? Can they feel? Are they autonomous agents? Or do they only imitate human minds, whose instruments they are, while they only *seem* to act like them?

Some answers to such questions from the field of cognitive science can be found in theses elaborated by graduate students of TIDD, among whose topics you can find “Artificial intelligence and bioevolution” (Alexandre Quaresma⁵), “Videogames as complex semiotic machines” (Henrique Bittencourt⁶), “Semiotics of artificial life” (Eduardo Camargo⁷), or “Electronic computers as autonomous semiotic agents” (Ricardo Gazoni⁸). Others can be found in several of the thematic issues of the electronic journal of TIDD, whose title is *Revista Digital de Tecnologias Cognitivas* (TECCOGS⁹). Among its issues that have dealt with the topic of cognition in some way or other are those on “collective intelligence”, “cognition and information”, “cognitive artifacts for building knowledge”, “artificial intelligence”, or “digital humanities”.

Quite a number of studies on issues of cognition conducted at TIDD have taken inspiration in the philosophy of cognition of Charles S. Peirce. This is why we are happy to have with us, today, the internationally renowned philosopher Vincent Colapietro, who is as much expert in Charles S. Peirce as he is on the philosophy of mind, having been translated into Portuguese with his book *Peirce e a abordagem do self*.¹⁰

5 QUARESMA DE MOURA, Alexandre Miranda. *Inteligências artificiais e bioevolução*. Tese (Doutorado), 2020.

6 BITTENCOURT NETO, Levy Henrique. *Videogames como máquinas semióticas complexas: a emergência dos interpretantes nos jogos digitais*. Tese (Doutorado), 2019.

7 CAMARGO, Carlos Eduardo Pires de. *Semiótica da vida artificial*. Tese (Doutorado), 2018.

8 GAZONI, Ricardo Maciel. *Semiótica da programação: levantamento crítico e perspectivas peircianas*. Tese (Doutorado), 2015.

9 <pucsp.br/pos/tidd/teccogs/>.

10 COLAPIETRO, Vincent. *Peirce e a abordagem do self: uma perspectiva semiótica sobre a subjetividade humana*, trad. Newton Milanez. São Paulo: Intermeios, 2014.

Professor Vincent Colapietro was Liberal Arts Research Professor of Philosophy at Penn State University until 2015 and is currently Professor in New England at the Center for the Humanities of the University of Rhode Island. He is well known in São Paulo. For decades, he has been a guest of the Catholic University at conferences organized by Lucia Santaela and Ivo Ibri, among others.

Why should Peirce, who lived from 1839 until 1914, have answers to questions of the philosophy of mind in the 21st century? This could be the leading question in our discussion with professor Colapietro today.

Peirce has actually written on machines in relation to minds, but his relevance to the contemporary philosophy of mind seems to go much beyond the paper of 1887 whose title was “Logical Machines”.¹¹ Peirce was the author of a philosophy that postulated the continuity between mind and matter, and hence human thinking and mechanical computing or reasoning. More specifically, he defended theses such as,

1. Every thought, or cognitive representation, is of the nature of sign.¹²
2. We cannot think without signs since all cognition, all thought, is in signs.¹³
3. The thoughts of a living writer are as much in any printed copy of his writings – and we may add today, in his technology of intelligence – than they are in his brain.¹⁴

¹¹ Logical machines, *American Journal of Psychology* 1 (Nov 1887), p. 165-170. W6: 65-72. – Port. 2014. CESTARI, Guilherme Henrique de Oliveira; GAZONI, Ricardo Maciel; NÖTH, Winfried. Tradução comentada de “Máquinas lógicas” de Charles S. Peirce. *TECCOGS: Revista digital de tecnologias cognitivas*, São Paulo, n. 10, p. 20-47. Disponível em: <pucsp.br/pos/tidd/teccogs/dossies/2014/edicao_10/1-dossie_1traducao_comentada_maquinas_logicas_peirce_guilherme_ricardo_winfried.pdf>. Acesso em: 23 set. 2020.

¹² On pragmatism: From a review of a book on cosmology. *CP* 8.191, 1904. [*CP* = PEIRCE, C. S. *Collected Papers*, vols. 1-6, HARTSHORNE, C. & WEISS, P. (eds.); vols. 7-8, BURKS, A. (ed.) Cambridge, MA: Harvard University Press, 1931-58. – References: *CP* followed by volume, full stop, paragraph number.]

¹³ Prolegomena to an apology or pragmatism: “There cannot be thought without signs” (*CP* 4.551, 1906).

¹⁴ Psychognosy: “It is much more true that the thoughts of a living writer are in any printed copy of his book than that they are in his brain” (*CP* 7.364, 1902).

4. Thought is not necessarily connected with a brain. It appears in the work of bees, of crystals, and throughout the purely physical world.¹⁵
5. It is a mistake to conceive of mind and matter as absolutely distinct.¹⁶

Are such ideas still relevant to the current debate on minds and machines? Dear Vincent, please feel free to choose from these questions and topics the ones that you find of greatest interest to the study of cognition at TIDD, and do not hesitate to address issues that you might find even more relevant. But if you don't mind, please tell us first about some of the links between your own career and studies in cognitive semiotics at the Catholic University of São Paulo.

v.c.: I must begin with a confession. As much as I have read in this transdisciplinary field, and I have read a bit, I am uncertain just what cognitive semiotics is. I know what it might be, and I know what such contemporary representatives of "Cogsem" as Jordan Zlatev take it to be.¹⁷ I also know what a Peircean would take it to be, but when I scan the field, I find that things sometimes need clarification. Imagine looking through a microscope and the thing at which you are looking comes in focus and then gets out of focus.¹⁸ On the prevalent assumptions of many of the leading contributors, the object of inquiry does not stay in clear and stable focus. I think of cognitive semiotics in that fashion: there seems to be an object of study that is finely focused at one moment and then there is a loss of focus a short time later. I find this principally because the relationship with Semiotics is inadequately thought through. Much of what calls itself cognitive semiotics, or for short, Cogsem, seems to have grown out of cognitive science. As a result, it seems not to have sufficiently deep roots in any semiotic theorist, including Peirce, so I must say I am a little bit confused when I read some of this. Is *Cogsem* fully or deeply enough semiotic?

15 Prolegomena to an apology or pragmaticism (CP 4.551, 1906).

16 Man's glassy essence (CP 6.268, 1891).

17 ZLATEV, Jordan. Cognitive semiotics: An emerging field for the transdisciplinary study of meaning. *The Public Journal of Semiotics* vol. 4, no. 1 (2012), p. 2-24.

18 Jordan ZLATEV's paper (see previous fn.) illustrates this point quite well.

If we approach it from Peirce's point of view, let us imagine what he might have said. Those papers from the late 1860s are part of what is called the "Cognition Series"¹⁹, as you well know. As such, they in effect inaugurate a program of cognitive semiotics. Cognition is, and your quotations sent to me in preparation for this exchange make this point very clearly, very solidly, essentially a semiotic process. All thinking and knowing depend on signs. A facet of this that needs to be highlighted because, with Peirce, the question of knowledge needs to be reframed. There is a sense in which Peirce does not have an epistemology – I know this sounds heretical and I suspect my listeners might be shocked to hear such a claim since Peirce wrote so much about knowledge. But terms have to be used very carefully, including the word *epistemology*. Peirce has no epistemology if we think of epistemology in its traditional form, the theory of knowledge crystallized around the question of skepticism. Epistemology in this sense was the attempt to either prove or disprove the very possibility of us knowing anything at all. To a great extent, Peirce was not interested in that question. He was not out to refute skepticism; rather he was committed to facilitating the growth of knowledge. He was not troubled, as contemporary epistemologists are, by whether knowledge is possible. Peirce was, as just noted, much more concerned with the *growth of knowledge*. There is thus with Peirce a shift – and I am following Joseph Ransdell²⁰ here – a shift *from* what, traditionally, epistemology has been *to* what methodology or methodetic ideally ought to be. For Peirce it is primarily a question of "how". – *How* can we take what we know or what we suppose we know and use that to discover what we do not know? The focus is upon learning and upon discovery. It is upon an open-ended process in which the very foundations of what we take to be secure might turn out, in the course of inquiry, to be insecure or unstable.

¹⁹ PEIRCE'S "Cognition Series" consists of following three articles, published in 1868 and 1869 in the *Journal of Speculative Philosophy*: "Questions concerning certain faculties claimed for man" (1868), "Some consequences of four incapacities" (1868), and "Grounds of validity of the laws of logic: Further consequences of four incapacities" (1869). (They have been republished in *CP* 5, 213-357, *EP* 1: 11-82 and *w* 2). In his article "The Journal of Speculative Philosophy Papers", C. F. DELANY provides a succinct but very helpful account of this series (in *w* 2, p. xxxvi-xlii).

EP 1 = *Essential Peirce*, vol. 1. HOUSER, N, & KLOESEL, C. (eds.). Bloomington, IN: Indiana University Press.

W 1, *w* 2, *w* 6 = *Writings of C. S. Peirce. A chronological edition*, vol. 1, 2, 6. MOORE, E. C. *et al.*, eds. Bloomington, IN: Indiana University Press, 1982, 1984, 2000.

²⁰ For example, RANDELL, Joseph. Some leading ideas of Peirce's semiotic. *Semiotica*, vol. 19 (1977), p. 157-178. RANDELL, Joseph. Charles S. Peirce. In SEBEOK, Thomas A. ed. *Encyclopedic Dictionary of Semiotics*, vol. 2, p. 673-695. Berlin: Mouton de Gruyter, 1986. Ransdell most emphatically made this point regarding methodology as distinct from epistemology in his Presidential Address to the Charles S. Peirce Society, "Peirce and the Socratic tradition". *The Transactions of the Charles S. Peirce Society*, vol. 36, no. 3 (2002), p. 341-56.

I read Peirce as an anti-foundationalist, but equally as an anti-skeptic. We do not need foundations in order to have knowledge. That is the Cartesian trap. to suppose that we need absolutely secure foundations. In supposing, we need such foundations, indeed, in identifying knowledge with absolutely certain truth, we are caught in one of the various Cartesian traps. What we need are self-corrective processes and procedures, what we have to institute, as best we can in the course of inquiry, are self-corrective processes and methods whereby we can detect and correct our errors. For me, then, the pursuit of knowledge is an ever more refined set of self-corrective practices. The emphasis then falls upon process, upon practice, upon embodiment, and upon what is called “enactment”. I think Peirce anticipated very clearly an extended theory of mind, so the question of the locus of mind is not in the individual consciousness. It is certainly not in an individual disembodied or purely spiritual consciousness.

Tools are extensions of the body, even such simple technologies as a hammer, a nail, a pencil, and a sheet of paper. It is illuminating to recall here a comment made in 1887 about a logical machine. “It has been contrived to do a certain thing, and it can do nothing else. For instance, the logical machines that have thus far been devised can deal with but a limited number of different letters. The unaided human mind is also limited in this as in other respects; but the mind working with pencil and plenty of paper has no such limitation. It presses on and on, and whatever limits can be assigned to its capacity today, may be over-stepped tomorrow. This is what makes algebra the best of all instruments of thought; nothing is too complicated for it.”²¹ But may not logical machines also overstep their present limitations? That is, after all, one of the most important if vexing questions we can pose regarding this matter. Imagine what a computer is, as a somatic extension. We pour our intelligence into these collective extended repertoires of artifacts. For me, following Peirce, mind is multiply located it does not have a single or absolute locus. There is a sense in which the community is the locus of intelligence, and my individual intelligence is realized first and foremost by means of my participation in the community of inquiry. The community of inquiry is not to be construed in any mystical or some overly ontological manner, but the community of inquiry is part of the very community of being. “All communication from mind to mind is,” Peirce insists, “through continuity of being”.²² My being is continuous with the very being of other beings. This alone makes knowledge possible. While Descartes begins with the assumption that consciousness is cut off from the world, also that it is cut off from

²¹ Logical machines (see fn. 10) and w 6, p. 70-71.

²² Immortality in the light of synechism (CP 7.572, c.1892).

its own body, Peirce begins with altogether different assumptions. Moreover, Cartesianism assumes the existence of language without explaining it. It also assumes the existence of an intelligence that is either pre- or extra-linguistic.

Whereas Descartes assumes all of that, Aristotle, Giambattista Vico, Hegel, and Peirce stand in a very different lineage. The body is essential. Other beings, and especially other selves, are essential. I can come to my self-understanding only in and through my relationship to others. This is what Peirce calls *tuism*, from the Latin familiar form of “you”, “tu”.²³ I am always addressing another whether consciously or with the intention to do so or not. Thought is always addressed to a “thou”, to a “tu”, to a “you”. Rather than a monological conception of thought, we have an irreducibly dialogical conception of thought where the I and the you are equally primordial. No I without a you. No you without an I. That means *we*, you and I together. So, “we” has a reality in Peirce it has in very few philosophers. It certainly had it in Aristotle. It certainly had it in Hegel. But these thinkers are the exception, not the rule.

My first run-through is that there is something exciting happening in what traditionally calls itself cognitive philosophy or cognitive science, or what in contemporary language calls itself cognitive semiotics. But there has not been a careful enough effort to think through the relationship between semiotics and the study of cognition, and cognitive semiotics has been unduly constrained by certain assumptions that we can leave signs aside and focus on meaning. We can think about thinking as computation,²⁴ and all of these presuppositions seem to me to be erroneous views. If we thought through cognitive semiotics in more strictly, carefully Peircean terms, we would not have to reinvent the wheel to the extent that folks in cognitive semiotics are reinventing the wheel. It seems to me that this would have to begin with a highly flexible but nuanced conception of semiosis – sign activity. From there, we should try to give it specificity and determination by context. Peirce’s notion of semiosis is in one sense very abstract and very general, but that enables it to be very concrete, because it can be applied to so many different contexts.

²³ See FISCH, Max H. Introduction. *w 1*, p. xxix, where Fisch quotes Peirce’s definition of *tuism* for the *Century Dictionary*.

²⁴ Quite in contrast to Peirce, who would have criticized the view of thinking as computation as reductionist, Kalevi KULL adopts this view in “Virus semiosis”. *Transobjeto: Grupo de estudo dos confrontos entre o realismo especulativo e o realismo peirciano*, June 29, 2020. Disponível em: <transobjeto.wordpress.com/2020/06/29/virus-semiosis>. Acesso em: 23 set. 2020.

I think your students are very lucky because you are coming to the study of cognition with a very powerful model of semiosis, mind, and cognition in all of its forms: from perception to the most abstract form of reasoning. You and Lucia Santaella are coming at these phenomena with a very powerful model of at least three things: of meaning, of mind, and of knowledge in all of its forms, but you are not doing it in an antiquarian or a purely scholarly way. You are attuned to the cutting edge of what is going on in the new technologies, of what are the implications of these new technologies and you are trying to marry the older with the newest – not merely the newer – the newest. What is the cutting edge and so coming at it from traditional disciplines like such as linguistics, literary studies, philosophy, psychology, including psychoanalysis. You are coming at these questions from traditional disciplines but with an attunement to what is just now emerging at the cutting edge of cultural processes and practices. I find that immensely exciting. And just one other thing and then I want to hear from you. You mentioned the various groups that make up the constituents of your audience, from design, education, social networks. One of the things that is necessary is to give individuals coming from these diverse backgrounds with these diverse interests a language where they could begin to talk about their common concerns, and what it seems to me Peirce does is precisely that. He gives folks with different interests from different backgrounds with different aspirations a common language in which they can communicate. We cannot learn from one another if we are speaking a completely foreign tongue. We need to have a common tongue, and then we can see the analogies between, for example, design and education.

There is something definitive about Peirce's emphasis on self-criticism. His theory is ultimately a critically reflexive theory. It is a critique of itself in its various applications, and it seems to me that part of the power of Peirce is that he gives us the resources to become ever more self-critical, self-conscious, and self-controlling. I find it immensely important to share a common language with people from disparate and different backgrounds that enable them to be critical of others but much more importantly to be critical of themselves.

w.n.: Thank you for these many ideas. May I pick up some of them that address questions such as artificial intelligence and machines more generally? I would like to hear more about some of your topics. The first was self-correction. You gave much emphasis to self-correction in humans but, as we know, self-correction is a feature precisely of artificial intelligence. Is this one of the points where we can see a convergence of human minds and artificially intelligent machines? I leave it with this question to address the other ones later.

v.c.: Thank you. This is a wonderful question. Let me back up a step and say that there is a very famous passage contained in volume five of Peirce's *Collected Papers*. It is where he distinguishes various grades of self-control. He begins by noting, "of course there are inhibitions and coördinations that entirely escape consciousness".²⁵ The passage drives toward the conclusion: "The brutes are certainly capable of more than one grade of control; but it seems to me that our superiority to them is more due to our greater number of grades of self-control than it is to our versatility".²⁶ At this point in the text, the speaker ("the Pragmaticist") is challenged: Is it not due to our facility of language?"²⁷ The Pragmaticist (presumably Peirce) responds by asserting: "To my thinking that faculty is itself a phenomenon of self-control."²⁸ Of course, Peirce is post-Darwinian. Whereas classical Western thought often draws an absolutely sharp line between the human being and other animals, for Peirce, as a post-Darwinian, the human animal is just that. The human being is the human animal, and there is no absolutely sharp line between the human animal and other animals, especially as to the question of self-control and self-consciousness. Descartes did not even think other animals were conscious, let alone self-conscious. There is some evidence that some animals other than humans are self-conscious. They have some awareness. They recognize themselves in the mirror. So we need to discuss these matters in fine detail and in close connection with the most recent empirical studies. There is no absolutely sharp line between the human animal and other animals, in particular, concerning self-control and self-consciousness. When Peirce defines scientific or experimental intelligence, he defines it as the capacity to learn from experience,²⁹ which is, in part, the capacity of experience to remake our habits such that they are more finely and fully attuned to the habits of the beings we encounter in experience.

To get to your question, I think Peirce blurs the line not merely between the human animal and other animals, but also between human intelligence and artificial intelligence. As you say in one of your articles that I just reread recently³⁰, the notion of "quasi" is very instructive and useful. He sometimes will talk about it "as if", for example, "as if there

25 Pragmaticism [*Prag.* 4] (*CP* 5.533-34 c. 1905).

26 *Ibid.*

27 *Ibid.*

28 *Ibid.*

29 See, e.g., *CP* 2.227, c.1897.

30 NÖTH, W. Cognição como semiose: semiótica cognitiva e Cogsem [Cognition as semiosis: Cognitive semiotics and Cogsem]. To appear in *Revista DeSignis* 34 (2021) [= *Semióticas cognitivas: Nuevos paradigmas*].

is an interpreter” or “as if there is an utterer”, and it seems to me that “quasi-mind” is an apt expression he uses. Peirce is hedging his bet. He doesn’t want to make too strong a claim, but he doesn’t certainly want to make too sharp a divide. This is his position with regard to the question whether machines are intelligent. In certain respects, they are more intelligent than human beings, for example, in terms of chess games. In terms of certain other kinds of activities, such as simply computing, they can far outstrip human intelligence. So, are they intelligent? Yes. Are they intelligent in some other way in the same way as human beings are intelligent? Yes, including questions such as, Can they learn? They can correct themselves. Can they correct themselves? –, which goes back to a point you made earlier. In terms of perception, there is no question that we can make machines that are really skillful at pattern recognition. In that, as far as I know, if they make mistakes in the recognition of patterns, they build those mistakes into refining the operations going forward. On the surface, it looks like these are quasi-intelligent, or let us just say, they are intelligent. Are they intelligent to the same degree as human beings? There is where I hesitate. There is where I hold out because I think there is an open-ended level of self-correction and self-control, and I am not sure that we are at that point with any machine we have been able to make.

w.n.: Thank you, I agree with all you said, but let me come to your other remarks concerning human nature. Peirce has certainly answers as to both the affinities between intelligent machines and humans and the differences. Let us only consider his paper on the “Neglected Argument for the Reality of God”.³¹ Of course, we do not expect intelligent machines to reflect on God, but there is a connection with another of the topics you addresses – the one of “dialogue”. You said that thinking is dialogical, but if thinking is dialogical, how could intelligent machines enter into dialogue with themselves? Now, with whom would an intelligent machine conduct a self-dialogue? On the other hand, is there not some element of dialogicity when a machine corrects itself? Does not self-correction presuppose some sort of self-dialogue? If you consider how computers learns to dominate complex games, such as chess or go, you discover that they advance in their competence as they play against themselves. Is that not a kind of dialogue as it is needed for a thinking being?

³¹ PEIRCE, C. S. A neglected argument for the reality of God, *CP* 6.452–480 and *EP* 2:434–446 (1908).

v.c.: This is a very good and equally challenging question. As you were speaking, I was thinking that machines can also be self-monitoring. They can detect something in their operating system that is not going well, when this is the case. However, let me shift the ground a little. As you were speaking, I was also thinking about another aspect of our topic, the affective dimension of human minds, consciousness, and the human psyche. It is not merely the emotional, but also the deeply affective relationship that constitutes the human mind. Is it possible for a machine to *care* about a human being – for another as other? Now, the relevance of this question is that such solicitude might care for you as the unique you that you are. Might not such solicitude be a condition for certain forms of dialogue? It is rather significant and underexplored that Peirce identified what traditionally were the three theological virtues, Faith, Hope, and Charity, as logical sentiments³², and he thought that one could not be fully logical unless one was animated. He redefines these terms that are both continuous with traditional understanding, but also quite innovative in their own right. One could not be fully logical without possessing these virtues, these sentiments, without being animated by faith, hope, and charity, and it seems to me that there might be truth there. Then, the limits of machine intelligence would be the limits of their affective commitments, or their ability to obtain and refine the so-called logical sentiments of faith, hope, and charity. Can they be intelligence in certain ways? Yes, might those limitations trace their route to their inability in the final analysis to love? Because, for Peirce, Agape³³, the capacity to love the other as the other is absolutely crucial. Here is one of those places where Peirce is unwilling to divide the cognitive and the affective. In him, they are together. The limits of artificial intelligence might be the limits of

32 PEIRCE, C. S. The doctrine of chances (1878): “It may seem strange that I should put forward three sentiments, namely, interest in an indefinite community, recognition of the possibility of this interest being made supreme, and hope in the unlimited continuance of intellectual activity, as indispensable requirements of logic. Yet, when we consider that logic depends on a mere struggle to escape doubt, which, as it terminates in action, must begin in emotion, and that, furthermore, the only cause of our planting ourselves on reason is that other methods of escaping doubt fail on account of the social impulse, why should we wonder to find social sentiment presupposed in reasoning? As for the other two sentiments which I find necessary, they are so only as supports and accessories of that. It interests me to notice that these three sentiments seem to be pretty much the same as that famous trio of Charity, Faith, and Hope, which, in the estimation of St. Paul, are the finest and greatest of spiritual gifts. Neither Old nor New Testament is a textbook of the logic of science, but the latter is certainly the highest existing authority in regard to the dispositions of heart which a man ought to have.” CP 2.655 and EP I, p. 150.

33 In Evolutionary love (1893), Peirce introduced the terms “Agapism” and “Agapastic evolution” for the view that creative love is operative in the cosmos. “The good result is here brought to pass, first, by the bestowal of spontaneous energy by the parent upon the offspring, and, second, by the disposition of the latter to catch the general idea of those about it and thus to subserve the general purpose” (CP 6.303).

genuine affection, of the logical sentiments. Can computers think? Unquestionably. Can they think in sophisticated, self-corrective, self-monitoring ways? Unquestionably. It is not so much falling short, but are they different from human intelligence in this regard? Yes.

One other point: to what extent can computers make wild errors? To what extent can they fall into mistakes that seem to call much too much into question? For Peirce, humans are deeply fallible animals, and this is, actually, why we have once managed to survive. – Well, we will see what the upcoming decades bring, but let us leave that to the side. Once we were able to survive, we turned that vulnerability into an immense advantage. We turned the vulnerability of fallibility into a tremendous advantage because all science is making mistakes intelligently. Fallibilism is not only our tendency to fail; it also means our ability to see that we have made a mistake then find, ascertain, the source, the root of the error and, then, correct it.³⁴ There is a continual process of self-correction, predicated on our ineradicable fallibility. This is actually a good thing, that we were so prone to making mistakes. “It is a truth well worthy of rumination that all the intellectual development of man rests upon the circumstance that all our action is subject to error. *Errare est humanum* is of all commonplaces the most familiar. Inanimate things do not err at all; and the lower animals very little”.³⁵ Now, can you make a machine to make as many mistakes as human beings do?

w.n.: You are perfectly right. People always speak of the mistakes computers and robots make. They never talk about the mistakes human make in the same situations. This is why we need the philosophy of cognition to answer such questions. Philosophers of mind have given diverse answers without having reached consensus. Some have focused on intentionality, others on consciousness: computers have none, humans have. Can you say something about these two topics?

34 In 1897, William JAMES, in a book dedicated to Charles Peirce (*The Will to Believe and Other Essays*. Cambridge, MA: Harvard University Press, 1979, 25), wrote: “Our errors are surely not such awfully solemn things. In a world where we are certain to incur in spite of all our caution, a certain lightness of heart seems healthier rather than this excessive nervousness in their behalf. At any rate, the fittest thing for the empiricist philosopher.” – That is, the empiricist philosopher cannot help but be a thoroughgoing *fallibilist*. A “contrite” fallibilist (a thinker willing to confess, “I was wrong”) can possess “a high faith in the reality of human knowledge” (CP 1.14, c.1896), arguably, only such a fallibilist can possess such a faith.

35 Detached ideas on vitally important topics, (CP 6.86, 1898).

v.c.: Right, so here is one of the places where it seems to me Peirce's notion of synechism is especially relevant. Of course, this is his doctrine of continuity and so there is no absolutely sharp line to be drawn here. Let us begin with intentionality. If we go back to Franz Brentano, who was, of course, the one who, drawing upon the Scholastics, introduced the notion of intentionality. He seemed to be committed to an anti-synechistic project. By means of the concept of intentionality, he wanted to determine the feature that differentiates mental from all other phenomena, distinguishing the mental from the physical, as though there were an absolutely sharp line of demarcation between the two. He also began in a somewhat introspective way, a mind reflecting upon its own activity. I think Thomas Short³⁶ and others have suggested something that is really crucial: that purposive behavior is the key out of which ever more refined and abstract forms of intentionality have evolved. No doubt, an animal is conscious in the sense that it has thoughts about something. Hence, it is just part of the very comportment or conduct of the animal to be oriented toward the world and to be animated by quasi purposes. I think what we understand by the intentional and intentionality in that narrow mentalistic sense can be explained naturalistically if we think about animal behavior. If we think about the animal looking for food, it takes this smell to be indicative of the presence of food nearby, but in fact, it may be mistaken about this. Nonetheless, it is it is searching for something. An object of desire is implicated in the activity of searching. There is an aboutness there. Hence, we do not have to posit any anything mysterious under the rubric of intentionality. This may not be altogether fully adequate, but I think it points us in the right direction. A naturalistic account can begin to render intentionality intelligible. It does so by taking conduct itself to be purposive or at least quasi-purposive.

As to the other question, the one regarding consciousness, it seems to me that Peirce is immensely useful on this. He spilled a lot of ink in his correspondence with William James, who obviously was very important in the history of Psychology, delineating "consciousness" in a very detailed way. Some of those letters are virtual essays, where Peirce is writing to James and talking about what consciousness might mean.³⁷ It seems to me that here is where the Peircean categories, specifically, the power of the categories as heuristic prompts and guides, just shines forth, because

³⁶ SHORT, Thomas. Semeiosis and intentionality. *Transactions of the Charles Sanders Peirce Society*, vol. 17, no. 3 (1981), 197-223 and SHORT, T., *Peirce's Theory of Signs* (Cambridge: Cambridge University Press, 2007), p. 175-77.

³⁷ See especially CP 8.249-315 (1897-1909).

Peirce distinguishes between monadic, sort of dream consciousness, polar or dyadic consciousness, and then synthetic or mediational consciousness. Much of contemporary cognitive science including cognitive semiotics wants to operate with a purely univocal notion of consciousness, but obviously, consciousness is a word possessing many meanings. It is anything but univocal, possessing at least three fundamentally different meanings: qualitative consciousness, polar consciousness, synthetic or mediational consciousness. The interesting thing is that the higher the form of consciousness, the more easily replicable it is by machines. The so-called lower forms of consciousness, purely qualitative consciousness, are the ones that are most difficult to replicate in machines. Peirce's philosophy and Mark Champaign's book³⁸ are rather good here. It would be very instructive to trace in detail Consciousness in reference to Peirce, or, more accurately, to the classification of the sciences (e.g., what consciousness means in the context of Phenomenology, or in those of the three Normative Sciences, etc.). – I think, Peirce has much to say about the question of consciousness. However, to take it up in a truly Peircean manner, we would have to be attentive to his emphases and classifications. What philosophers ought to do well, Peirce does exceedingly well and that is to draw the relevant distinctions and not to suppose that we can get along with one word or with a univocal sense of such a complex nuanced word as *consciousness*. There are tremendous resources in Peirce's phenomenology and his theory of signs. Even his cosmology addresses questions of consciousness. I also think that there are resources for understanding intentionality here.

w.n.: We are coming to more and more difficult topics. Before we open the floor for questions, let me propose a no less difficult question. Life, for example, does intelligence presuppose life? How about artificial life, is life a dividing line between machines and biological beings or not?

v.c.: I tend to think that life is a condition for intelligence since the problem is: nothing is alive except in an environment, and nothing could sustain itself apart from an ambience. So, machines have a history that suggests a kind of life, that is to say, they have obviously evolved in conjunction with natural intelligence. However, natural intelligence has just as obviously evolved because of the ability to use and refine ever-new technologies. Human intelligence is dependent upon artificial extensions, but

³⁸ CHAMPAGNE, Mark. *Consciousness and the philosophy of signs: How Peircean semiotics combines phenomenal qualia and practical effects*. Cham: Springer, 2018.

consider, in Peirce's early writings, that little dialogue between human beings and the words created by them.³⁹ The words created by humans seem to be utterly dependent upon their creators, but Peirce argues that this is not entirely the case. There is a mutual dependence between language and humans. The words could also turn around and say, "You can think what you think because of us". It might very well be that this dialogue is an analog of understanding the relationship between naturally intelligent beings and artificially intelligent ones.

w.n.: Indeed, this seems to be the very key to the question of artificial and natural intelligence. Perhaps you can explain the quote to which you only made a very brief reference a bit more to our audience.

v.c.: What Peirce says is that words only mean what we make them mean, but then he imagines that these words may turn around and say, "That is not the whole story, nor the complete picture. You, human beings also mean only what we allow you to mean". Thus, there is this deep mutual dependence between language and humans. Language obviously depends upon us having somehow created it, and somehow, we have refined it over decades and centuries, but not adequately for many of our purposes. However, in turn we also depend utterly upon language as human beings. This is a case of mutual dependency – we depend on language as much as it depends on us.

w.n.: Right and we do not only depend on language, but also, more and more, on intelligent machines. This brings us to the end of our dialogue, but not to the end of our program because we have questions from the audience. Luis Felipe⁴⁰ will tell us what these questions are. Pollyana Ferrari⁴¹: "Isn't it the ability to love that makes us human? Would the machine's limit be affection?"

v.c.: I think so. I would say that this ability is only one among numerous abilities; I think our ability to make fantastic and wild mistakes is also very distinctive of the human animal. Even so, I do not want to

39 PEIRCE, C. S. Some consequences of four incapacities (1869): "Does not electricity mean more now than it did in the days of Franklin? Man makes the word, and the word means nothing which the man has not made it mean, and that only to some man. But since man can think only by means of words or other external symbols, these might turn round and say: 'You mean nothing which we have not taught you, and then only so far as you address some word as the interpretant of your thought.' In fact, therefore, men and words reciprocally educate each other; each increase of a man's information involves and is involved by, a corresponding increase of a word's information." (CP 5.313)

40 Thanks are due to Luis Felipe Napoli, student of TIDD, who managed the interface of this dialogue.

41 Pollyana Ferrari is a professor of TIDD.

identify us simply or solely with our ability to love. It is *a* defining feature of the human animal, but I resist making it *the* definitive property of human beings. It is a form of transcendence. To transcend the self in a way that allows for transformations of the self. Peirce was somehow drawn to Buddhism in this way and to certain forms of Christianity in this regard as well. However, I do not want to say that our ability to love is what marks us off from all other beings. I suppose other beings have the capacity to love, too, but I do think if machines are limited in their affection, they are limited in their intelligence. I do think that.

w.n.: Second question, Peter: “Can AI produce something like a soliloquy like Hamlet, who asked “To be, or not to be, that is the question?”

v.c.: Well, there are actually two questions there. There is a question about a soliloquy in general, and then there is a question specifically about a soliloquy in which one holds up the possibility of exterminating one’s own life. Those are two different questions, related, obviously, but not necessarily the same. To the first I say yes, you made the point, and you were just dead right in making it. Higher forms of artificial intelligence can engage in self-dialogue and a soliloquy is just another name for a self-dialogue, right, the self in dialogue with itself. That is the first question. The other question has to do with whether the machine could raise the issue of self-extermination or self-annihilation, the question of suicide. There, I plead ignorance. I think at present there does not seem to be much evidence to support that, but who knows what might come in the future?

w.n.: Question number three, Henrique Bittencourt, “Could you comment on the role of fallibilism in cognition?”

v.c.: Yes. There are various ways of putting it, I would say. One way to put it is to insist that fallibility is the engine that drives the pursuit of knowledge, that is to say our fallibility is the engine we are caught up short, time and again, by making a mistake. It is precisely that, which animates us to actually engage in the arduous work of inquiry. John Locke says, and I think Peirce was aware of this passage and actually influenced by Locke: The love of truth is, in some ways, easy. The love of *finding* the truth is a difficult love.⁴² Everybody wants the truth handed to them, but the arduous, oftentimes humiliating work is finding the truth, be-

⁴² *Essay on Human Understanding*, Book IV, chapter XIX.i (“On Enthusiasm”): “Everyone in the commonwealth of learning professes himself to be a lover of truth, and every rational creature would be offended if it were thought that he is not. And yet it’s true to say that very few people love truth for its own sake, even among those who persuade themselves that they do. How can anyone know whether he is seriously a lover of truth? I think there is one unerring mark of it, namely that one doesn’t accept any proposition with greater assurance than is justified by the proofs one has for it.”

cause you make so many mistakes. Mistake after mistake, the arduous, sometimes humiliating, work of seeking the truth, that is a difficult love. I think there is no more important role in cognition because cognition is never knowing. It is always the pursuit of knowing, always *coming to know*. There is no more important feature than our fallibility. Our limitless growth is actually rooted in our ineradicable fallibility.

w.n.: Without pretending to reach a conclusion, our discussion leaves us space for this last question: “Is our cognitive development limited through language? In the case of AI, through a language made up of 0s and 1s (0/1)?” What kind of language is that? Don’t we deserve a better language?

v.c.: Right. We may use the word “language” in a very narrow sense to designate natural languages such as Portuguese, or English, or French, or Spanish. Or we may use this word in a much broader sense to mean any mode of symbolization. I think that the limits of artificial intelligence are in part defined by the limits of our capacities, our symbolic capacities, not necessarily the limits of our natural language. We are limited but, for Peirce, this is always only a provisional limit. There is no ultimate invincible limit to what human ingenuity or imagination might yet craft by means of symbols.

w.n.: Such as Kant’s unsurpassable limit of the of the *unrecognizable thing in itself*, which Peirce rejected⁴³, is that what you mean?

v.c.: Yes, it would be, and it seems also that Peirce is ultimately an anti-formalist. He understands the power of algorithms, but he also understands that not everything is reducible to algorithmic relationships. By an algorithm, I mean a finite set, usually a very small set, of explicit rules that can be rendered fully explicit. Now, for certain purposes in certain contexts, algorithms are immensely powerful tools, but there is not an algorithm for everything. There is a role for some norm-guided but not strictly rule-governed forms of human ingenuity. So, yes, artificial intelligence is limited by the symbols we are able to use and fall back on. Yes, a language of just “yes” and “no”, one and zero, is limiting in some ways, empowering in others, but there is also this issue of algorithms, and Peirce is not ultimately a formalist who thinks that everything is formalizable. It is very interesting that Peirce hardly says anything about

⁴³ “The *Ding an sich*, however, can neither be indicated nor found. Consequently, no proposition can refer to it, and nothing true or false can be predicated of it. Therefore, all references to it must be thrown out as meaningless surplusage” (CP 5.525, c. 1905).

codes, in the European tradition. For Ferdinand de Saussure and Roland Barthes, there is much emphasis on codes and little attention to habits. For Peirce, the opposite is the case: habits are primordial, whereas codes are, being codifications of habits, derivative. In Peirce's theory of signs, habits in effect replace codes. He can account for codes by means of the codification of habits? Can Saussure and his progeny however account for habits?

w.n.: This is where we have to conclude, not because we have reached a final conclusion, but because we have wrapped up as many topics of interest to the students of our program in the short time available. Your contribution to our program TIDD digital was most important. We are very grateful, dear Vincent. Your presentation and your answers show that philosophy is needed in our field of studies because it offers answers to so many open questions of topical interest. Let us stay in contact.

v.c.: Thank you, goodbye then.

Biopolítica digital e a utilização da inteligência artificial no controle de pandemias

Paola Cantarini¹

Resumo: A era da quarta revolução industrial, chamada indústria 4.0 ou era do silício, caracteriza-se, sobretudo, pela utilização da inteligência artificial em todos os setores da vida, com o aumento da intensidade de interconexões técnicas. Época da biopolítica digital, da hipermodernidade e do transumanismo, pode-se observar que, nesse contexto, surge uma nova forma de soberania, sobrepondo-se à soberania baseada na manutenção das fronteiras entre países. O soberano agora é aquele que dispõe de dados. O mais recente uso da inteligência artificial encontra-se no controle da pandemia do Coronavírus, para o cumprimento pela população de medidas de isolamento social, bem como para o monitoramento do surto. Isso traz questões relativas ao conceito de estado de exceção e a necessidade da aplicação do princípio da proporcionalidade quando surgem casos de colisões de normas de direitos fundamentais, como ocorre diante de tais situações excepcionais. Este artigo pretende refletir sobre essas questões.

Palavras-chave: Biopolítica. Inteligência artificial. Controle de pandemias.

¹ Mestre e doutora em direito pela PUC-SP, com estágio doutoral na Universidade do Minho-Pt., doutora em filosofia do direito pela Univ. do Salento-Lecce-It., pós-doutorado em Ciências Sociais – Universidade de Coimbra-Pt., pós doutorado em Filosofia, Artes e pensamento crítico -EGS- Suíça. Pós-doutoranda – TIDD-PUC-SP. Pesquisadora UNICAMP. Pesquisadora do *lawgorithm*, e do IEA – Instituto de Estudos Avançados da USP. *Visiting Researcher* SNS-PISA-It. cv Lattes: lattes.cnpq.br/5057397566570034. E-mail: paolacantarini@gmail.com.

Digital biopolitics and the use of artificial intelligence to control pandemics

Abstract: The era of the 4th industrial revolution, industry 4.0 or the silicon era is characterized, above all, by the use of artificial intelligence in every sector of our lives, with the increase of technical interconnections of all species, with the acceleration of time. These are the times of digital biopolitics, hypermodernity and transhumanism. A new form of sovereignty emerges, sovereignty based on maintaining borders between countries is overlapped. The sovereign is now the one who owns data. The most recent use of artificial intelligence refers to the control of the Coronavirus pandemic, regarding population's compliance with quarantine measures, as well as monitoring the outbreak, bringing questions related to the concept of state of exception and the need to apply the principle of proportionality when faced with cases of collisions of fundamental rights norms, as in the case of such exceptional situations. This paper intends to reflect on such issues.

Keywords: Biopolitics. Artificial intelligence. Pandemic control.

Introdução

Vivemos na fase da hiperhistória, significando a dependência de nosso bem-estar das tecnologias da informação e comunicação. Esta fase da sociedade da informação ou de dados, hoje considerados como o novo petróleo, dá origem a uma nova forma de soberania, a soberania digital, já que agora o soberano é aquele que dispõe de dados, nos termos Byung Chun-Han (2020). O extraordinário avanço tecnológico e da inteligência artificial, em todos os setores de nossas vidas, intensificou as interconexões técnicas, trazendo a percepção de aceleração do tempo.

A inteligência artificial (IA) traz diversas mudanças em todos os setores sociais e, certamente, no Direito, ocorrendo a virtualização do nosso eu, de nossas personas e corpos, a gestação de subjetividades em contínua mutação e novas concepções de tempo e de espaço. Acreditamos em imagens técnicas (FLUSSER, 2017) que se afastam dos conceitos e da realidade, sem enxergar tal distanciamento. Distanciamos-nos de nós mesmos, em busca da liberdade sem limites, mas não temos mais como pousar, presos que estamos na pós-história e em prisões virtuais e digitais. O corpo e a subjetividade, na sociedade digital pautada no controle microprotético e midiático-cibernético principalmente, são regulados por um conjunto de tecnologias biomoleculares, digitais e de transmissão de informação. Surgem subjetividades digitalizadas e cada vez mais performáticas e superficiais, pois a empatia vai se perdendo a cada geração, enquanto a linguagem escrita vai se transformando em oralidade cibernética. Trata-se da era do “phylum maquínico”, termo forjado por Gilles Deleuze e Felix Guattari, relacionando-se tal expressão ao silício no agenciamento contemporâneo homem-natureza, ou seja, uma linhagem tecnológica que emerge a cada vez que um conjunto de singularidades, prolongáveis por operações convergem e as fazem convergir para um ou vários traços de expressão assinaláveis (DELEUZE; GUATTARI, 1997, p. 87).

Nesse contexto a IA impacta na construção do Direito, envolvendo os temas relacionados ao Direito Público e a Teoria do Estado, Direito Constitucional, tais como soberania digital, estado de exceção, bem como o princípio da proporcionalidade, da razoabilidade e os direitos funda-

mentais. Somos dominados por conceitos alienantes e imagens técnicas que nos subjugam, pois nos fazem crer serem a realidade, quando, na verdade, se distanciam ainda mais dela, representando, assim como a abstração matemática, um fim em si mesmo, levando-nos a cair na armadilha do domínio técnico dos aparelhos, sem perceber. Edmund Husserl (2006), em sua doutrina do conceito, entende que a representação própria e direta de um objeto somente seria alcançada pela intuição, enquanto que o conceito limitar-se-ia a fornecer uma representação imprópria, simbólica ou mediante símbolos com um caráter intencional. Intencionar é tender por meio de conteúdos dados à consciência a outros conteúdos que não são dados. Nesse sentido, a produção de “conhecimento” por meio da IA, em uma forma de cálculo, traz a problemática da perigosa alienação na técnica trabalhada por Husserl em seus textos reunidos sob a rubrica “Crise [das ciências e] da humanidade europeia” (ibid.), do final de seu percurso filosófico, já que há a construção de um universo simbólico apartado das evidências da intuição sensível, a qual está impossibilitada de ser produzida por uma máquina. A ciência se descola do mundo da vida, do mundo vivido, da vivência mundana, e logo, de nós, seres humanos.

A ciência, saber característico da modernidade, valoriza apenas a causa formal, em detrimento das outras causas segundo a doutrina de Aristóteles, formalismo que fundamentou toda a ciência moderna bem como o domínio científico; desde a revolução copernicana, com Galileu Galilei, provocando o distanciamento dos fundamentos do saber antigo e medieval, de cunho aristotélico. Na modernidade, não distinguimos mais o que é forma e matéria, eficiência e finalidade, estando todas reduzidas a uma só causa, a causa formal que, acionada pela causa eficiente, dá acesso ao conhecimento, ao poder de dominar as coisas e, assim, influenciar o curso do universo. Antigamente, portanto, o objetivo da ciência não era um objetivo tão econômico e utilitário quanto se tornou após Newton, muito diverso, portanto, do que era postulado pela alquimia. Já na década de 1930, em seus estudos sobre a “Crise da civilização europeia”, antes referidos, Husserl, repensou as matrizes europeias sob o impacto da Grande Guerra, que se tornaria, depois da II, a I Guerra Mundial, condenando o modo ocidental de estudar a realidade, ressaltando a característica moderna de intervir na realidade. As condições denunciadas por Husserl intensificaram-se no ritmo de desenvolvimento das ciências no século XX até o ponto em que a humanidade hoje se encontra diante do desenvolvimento acelerado da IA e, ao mesmo tempo, profundamente atropelada pela Covid-19.

IA e controle da pandemia

O mais visado uso da IA refere-se ao controle da pandemia do Coronavírus, ao cumprimento pela população de medidas de isolamento social, bem como ao monitoramento do surto e à aceleração de testes de medicamentos envolvendo a área da vigilância global de doenças, com destaque para a empresa *Northeastern University*. Recuperando notícias conhecidas, em 31 de março de 2019, a empresa canadense *Blue-Dot*, utilizando-se de IA para revisar mídias e redes sociais, detectou a propagação de uma doença incomum na China, em Wuhan, antes das primeiras manifestações da ONU e da OMS. Diversos países, tais como China, Coreia do Sul e EUA conseguiram antecipar possíveis áreas de contaminação utilizando-se da IA (aplicativo *Private KIT*), que se utiliza da localização de *smartphones* para verificar se as pessoas estão seguindo as diretrizes governamentais no tocante à quarentena e ao isolamento social.² A Europa previu a utilização do rastreamento Pan-Europeu de proximidade e preservação da privacidade (PEPP-PT) por meio da utilização de *smartphones* para o monitoramento da propagação do vírus e o cumprimento de ordens de quarentena³.

No Brasil, por sua vez, a USP criou a plataforma digital “Coronabr” em que um enfermeiro virtual ajuda a identificar se a pessoa precisa ir a um posto de saúde ou não. Algumas *startups* utilizam inteligência artificial, por exemplo, um robô para aferir a temperatura corporal de pessoas com suspeita de coronavírus, com maior precisão do que se tal trabalho fosse realizado por um humano e com menores riscos de maior propagação do contágio.⁴

As operadoras de telecomunicação estão atuando em conjunto com o Ministério da Ciência, Tecnologia, Inovações e Comunicações (MCTIC), afirmando a preservação do anonimato dos clientes, sem, contudo, disponibilizar informações suficientes para se comprovar tal alegação, ferindo com isso o princípio constitucional da transparência e o princípio constitucional da publicidade. Rio de Janeiro e Recife divulgaram parcerias quanto ao mapeamento de dados de geolocalização, sendo destacado pela Prefeitura de Recife o compromisso com a “*privacy by design*”, ou seja, que os dados pessoais coletados serão mantidos anônimos, condicionando sua utilização ao consentimento livre, expresso e informado dos usuários.⁵

2 Disponível em: <correio braziliense.com.br/app/noticia/ciencia-e-saude/2020/03/26/interna_ciencia_saude,839865/inteligencia-artificial-pode-ajudar-na-prevencao-de-pandemias.shtml>. Acesso em: 20 maio 2020.

3 Disponível em: <cnnbrasil.com.br/tecnologia/2020/04/01/grupo-europeu-desenvolve-tecnologia-de-rastreamento-para-conter-covid-19>. Acesso em: 20 maio 2020.

4 Disponível em: <poder360.com.br/coronavirus/site-tem-enfermeiro-virtual-que-analisa-se-usuario-tem-coronavirus>. Acesso em: 20 maio 2020.

5 Disponível em: <g1.globo.com/pe/pernambuco/noticia/2020/03/24/reci-

O governo de São Paulo, utilizando-se de dados de celulares para fins de controle do cumprimento da quarentena, criou o sistema de monitoramento inteligente, denominado “SIMI-SP”, com a parceria de operadoras de telefone Vivo, Claro, Oi e Tim, mas sem qualquer consentimento ou permissão pelos usuários. Tal sistema permite o envio de mensagens de alerta para regiões com maior incidência da COVID-19; com o SIMI-SP o Governo poderá consultar informações georreferenciadas de mobilidade urbana em tempo real, e informa que, para garantir a privacidade, o monitoramento será coletado em conglomerados a partir de 30 mil pessoas.⁶

Em artigo publicado no site do Jornal Estadão⁷, Ricardo Campos, Juliana Abrusio e Juliano Maranhão destacam a diferença no uso de aplicativos utilizados no programa de monitoramento do cidadão desenvolvidos na Áustria e Alemanha, com destaque para os aplicativos dos Institutos Robert Koch (RKI) e Heinrich Hertz (HHI) da Alemanha em comparação com aqueles de Singapura e China, já que aqueles permitiriam que os dados pessoais permaneçam anônimos e encriptados, não sendo criados perfis de movimentos de cidadãos, possibilitando o mapeamento de cadeias de infectados decorrentes da proximidade com um infectado anterior. Em suas palavras: “não se monitoram indivíduos, apenas são mapeadas relações e contatos; [...] esta é a diferença fundamental para os aplicativos provenientes da Ásia, que também registram a localização de um utilizador e enviam todo o pacote de dados para as agências governamentais”. Os autores ressaltam, outrossim, a importância do respeito aos direitos fundamentais dos cidadãos, sob pena da crise sanitária se tornar uma crise democrática, e o prenúncio de um estado de exceção.

Com a utilização da IA no controle da pandemia do coronavírus, esta seria responsável pela escolha de quais casos seriam prioritários ou não para o uso de aparelhos de ventilação pulmonar. São, entretanto, questionáveis os métodos de escolha de qual ser humano, diante da impossibilidade de se atender a todos, deverá sobreviver, e quem deverá ser sacrificado, já que todos possuem a mesma dignidade humana e o direito humano à respiração para viver (MBEMBE, 2020). A crise sanitária é indissociável da discussão de questões eminentemente éticas, em especial no tocante à triagem e ao momento apropriado para a suspensão do confinamento

[fe-rastreia-700-mil-celulares-para-monitorar-isolamento-social-e-direcionar-a-coes-contra-coronavirus.ghtml](#)>. Acesso em: 20 maio 2020.

⁶ Disponível em: <tjsp.jus.br/Download/Portal/Coronavirus/Material/Cadip-Covid19.pdf?637354710131705136>. Acesso em: 20 maio 2020.

⁷ Disponível em: <politica.estadao.com.br/blogs/fausto-macedo/covid-19-co-mo-promover-a-saude-publica-e-protoger-a-privacidade>. Acesso em: 20 maio 2020.

social. Tal crise se relaciona com a crise ecológica e somente poderá ser resolvida com a realização de reformas estruturais imprescindíveis, como o aumento da proteção ambiental, a redução do aquecimento global, a proteção da Amazônia, a reorganização da economia global de forma independente dos mecanismos de mercado, a melhora da condição de vida da população em estado de vulnerabilidade, a redução do consumismo desenfreado e a criação de impostos realmente progressivos.

Habermas (2020), no artigo “Fios do tempo: precisamos agir como o saber explícito do nosso não saber”, ressalta a questão ética acerca da escolha trágica, a ser realizada pelo médico ou por meio da IA, entre uma vida e outra, surgindo a tentação de se violar o princípio da igualdade estrita de tratamento, favorecendo os mais jovens em detrimento dos mais idosos, afirmando que não podemos nos valer de fundamentos utilitaristas e lógicas utilitaristas de cálculo. Ressalta, outrossim, que a restrição de um grande número de importantes direitos à liberdade jamais poderá perdurar no tempo, devendo haver uma duração muito determinada. Preocupa ao autor, muito conhecido também por obras sobre Direito (HABERMAS, 1997), certamente, a falta de legislação para respaldar as medidas restritivas de direitos fundamentais, como a liberdade, apesar de necessárias, logo, proporcionais, e bem aceitas pela população, dando ensejo ao que o jurista alemão Carsten Bäcker (2020) qualificou como “estado de exceção discursivo”.

Nem sempre a opção pelo princípio da maior eficácia seria a melhor abordagem em se tratando de direitos humanos e fundamentais, como também em termos humanitários, sendo que, em um Estado Democrático de Direito, jamais um interesse econômico poderá se sobrepor a vidas humanas.

Consequências da utilização de IA no contexto da pandemia

Byung-Chul Han (2020) entende que os países asiáticos estão lidando melhor com tal crise do que o Ocidente, já que trabalham com dados e máscaras. Estados asiáticos como Japão, Coreia, China, Hong Kong, Taiwan e Singapura possuem uma mentalidade coletivista e respeitosa à autoridade, de inspiração confucionista, o que facilita a obediência pela população, apostando, sobretudo na vigilância digital por meio de *big data*, contudo, quase sem consciência crítica diante de tal vigilância. Na China não existe nenhum momento da vida social que não esteja submetido à observação, ocorrendo uma irrestrita troca de dados entre os fornecido-

res da Internet e de telefonia celular e as autoridades, e quase não se fala em proteção de dados. Há 200 milhões de câmeras com reconhecimento facial, dotadas de inteligência artificial por todas as ruas, capturando a imagem e medindo a temperatura corporal. Taiwan e Coreia seguem a mesma dinâmica da biopolítica digital, uma nova forma de soberania.

Diversos filósofos e pensadores apontam para a desproporção nos métodos utilizados até o momento, pois, mesmo em caso já confirmado de pandemia como a do coronavírus, o que se verifica, em especial, no caso da China e Índia, são notícias e imagens da utilização de violência física por parte dos policiais contra pessoas que circulavam pelas ruas deste último país. Neste sentido, Han (ibid.) destaca também a reação desproporcional e desmedida do pânico desatado pela nova pandemia, sendo esta uma reação imunitária social e global face ao novo inimigo, desdobramento daquela crise autoimunitária que já temos anunciado desde há tempos, como na tese de filosofia do direito que defendemos em 2017 na *Università del Salento* (Itália), intitulada “Princípio da proporcionalidade como resposta à crise autoimunitária do Direito”, em 2017 (CANTARINI, 2017).⁸ Outro motivo para o pânico é a digitalização, por eliminar a realidade, suprimindo a negatividade da resistência, surgindo uma apatia frente à realidade.

Na era da biopolítica digital, o conhecimento é definido como conhecimento para o mercado, sendo este o mecanismo principal de validação da verdade. Os mercados se transformam cada vez mais em estruturas e tecnologias algorítmicas, considerando como o único conhecimento útil o conhecimento algorítmico, daí a conclusão de Byung-Chul Han, de que agora o soberano é aquele que dispõe de (e sobre os) dados.

O estado de exceção, como é bem notório, foi trabalhado em diversos livros pelo filósofo italiano Giorgio Agamben, o qual em dois recentes artigos, “La invención de una epidemia” (2020a) e “Contagio” (2020b) associa tal questão à atual pandemia do coronavírus, por terem sido adotadas medidas excepcionais de emergência pelo governo italiano, trazendo também um alerta acerca da possível desproporcionalidade das medidas tomadas e afirmando a presença mais clara do que nunca de que vivemos em um estado de exceção generalizado. Em outro artigo, o mesmo pensa-

⁸ Destaca-se, da mesma forma, a investigação posteriormente desenvolvida como pesquisadora visitante na *Scuola Normale Superiore di Pisa-It.*, tendo por professor tutor, Roberto Esposito. Esposito ao lado de J. Derrida são dois filósofos que abordam a temática da imunidade, na filosofia e na religião, respectivamente, enquanto que Willis S. Guerra Filho trabalha com tal temática relacionada ao Direito.

dor afirma novamente: “É perigoso aceitar qualquer limitação da liberdade em nome da segurança, típico dos estados de exceção que não seriam estados de direito, mas um estado de controles cada vez mais generalizados” (AGAMBEN, 2015). Em tal estado, o pacto social se altera, cria-se um pânico generalizado e em face de tal situação de stress permanente as pessoas acabam cedendo facilmente sua liberdade sem qualquer limite.

A utilização da IA no combate a pandemias, bem como a determinação de quarentena e isolamento social de forma obrigatória envolvem necessariamente a análise do conflito de direitos fundamentais. Recente Resolução I-2020 da CIDH - Comissão Interamericana de Direitos Humanos denominada “Pandemia y derechos humanos em las Américas”, de 10 de abril de 2020, destaca o princípio da proporcionalidade como forma de evitar a generalização de um estado de exceção em tempos de pandemia, ressaltando sua limitação temporal e a obrigatoriedade de observância do princípio da proporcionalidade em seus três subprincípios, proporcionalidade em sentido estrito, adequação e necessidade (item 21). Também enfatiza que o estado de exceção cumpra e respeite os direitos humanos (item 20).

Ocorre aqui uma contradição ou paradoxo, pois o estado de exceção se caracteriza justamente por ser contrário ao Estado de direito, em que se respeitam os direitos humanos e direitos fundamentais, em especial por parte do Estado, ocorrendo o inverso no estado de exceção, como já bem demonstrou a história. O importante é nos assegurarmos da limitação no tempo de eventual concretização do estado de exceção e das limitações necessárias quando da suspensão ou restrição de direitos fundamentais, com base nos três subprincípios da proporcionalidade. Neste ponto é preciso ter em vista o princípio da proporcionalidade em sentido estrito, seguindo-se os desenvolvimentos neste sentido por Willis Santiago Guerra Filho (1989, p. 74-76), diverso de autores que se consagraram tratando do assunto, como Robert Alexy (2008) e Virgílio Afonso da Silva (2002, p. 26-50), desenvolvimentos que determinariam que tal limitação ou restrição jamais possa ocorrer, não importa o quanto resulte em benefício para outro princípio ou direito fundamental, a ponto de ferir o núcleo essencial de todo direito fundamental e, porque não dizer, de todo verdadeiro Estado Democrático de Direito, a saber, a dignidade humana.

Considerações finais

A utilização da IA no combate a pandemias, bem como a determinação de quarentena e isolamento social de forma obrigatória envolvem necessariamente a análise do conflito de direitos fundamentais. O fundamental é assegurarmos sempre, no caso de colisões de normas de direitos

fundamentais ou de princípios constitucionais, a observância do princípio da proporcionalidade, por meio da aplicação de um procedimento objetivo e racional, ressaltando a limitação temporal de restrições a direitos fundamentais diante de quadros de exceção, como no caso de pandemias.

Da mesma forma devem ser observados os princípios éticos e jurídicos relacionados à temática da utilização de IA, com destaque para as pautas éticas do *Ethics Guideline for Trustworthy AI* publicado pela Comissão Europeia em 8 de abril de 2019, as quais já foram positivadas em sua grande parte na forma de princípios jurídicos pela LGPD e pelo Regulamento Geral de proteção de dados da UE, sendo essencial a observância do respeito à autonomia humana, à explicabilidade, à transparência, à não discriminação, e à justiça. Também é essencial a observância dos requisitos para uma IA confiável com destaque para a intervenção e supervisão humana, ou seja, é fundamental o controle humano da tecnologia e a promoção dos valores humanos.

Referências

AGAMBEN, Giorgio. O estado de emergência não pode ser permanente. *La Repubblica*, 25 nov. 2015. Disponível em: <ihu.unisinos.br/169-noticias/noticias-2015/549436-giorgio-agamben-qo-estado-de-emergencia-nao-pode-ser-permanenteq>. Acesso em: 12 set. 2020.

_____. Contágio. In: AGAMBEN, Giorgio, et al. *Sopa de Wuhan: pensamiento contemporáneo en tiempo de pandemias*. [S.l.]: Editorial ASPO, 2020a, p. 31-33.

_____. La invención de una epidemia. In: AGAMBEN, Giorgio, et al. *Sopa de Wuhan: pensamiento contemporáneo en tiempo de pandemias*. [S.l.]: Editorial ASPO, 2020b, p. 17-19.

ALEXY, Robert. *Teoria dos direitos fundamentais*. São Paulo: Malheiros, 2008.

BÄCKER, Carsten. Corona in Karlsruhe: eine Prognose. *Verfassungsblog*, 25 mar. 2020. Disponível em: <verfassungsblog.de/corona-in-karlsruhe>. Acesso em: 12 set. 2020.

CAMPOS, Ricardo; ABRUSIO, Juliana; MARANHÃO, Juliano. Como promover a saúde pública e proteger a privacidade. *Estadão*, 09 abr. 2020. Disponível em: <politica.estadao.com.br/blogs/fausto-macedo/covid-19-como-promover-a-saude-publica-e-protetger-a-privacidade>. Acesso em: 12 set. 2020.

CANTARINI, Paola. *O princípio da proporcionalidade como resposta à crise autoimunitária do Direito*. Rio de Janeiro: Lumen Juris, 2017.

_____. *Elementos para um direito emancipatório e contra-hegemônico: uma análise crítica dos direitos humanos e fundamentais de grupos vulneráveis*. Rio de Janeiro: Lumen Juris, 2019.

DELEUZE, Gilles; GUATTARI, Felix. *Mil platôs*. São Paulo: Editora 34, 1997.

ESPOSITO, Roberto. *La Repubblica*, 21 abr. 2011; Disponível em: <ihu.unisinos.br/noticias/502044-filosofia-do-bem-comum-artigo-de-roberto-esposito>. Acesso em: 12 set. 2020.

FLUSSER, Vilém. *O mundo codificado: por uma filosofia do design e da comunicação*, São Paulo: Ubu, 2017.

GUERRA FILHO, Willis Santiago. *Ensaio de teoria constitucional*. Fortaleza: Imprensa Universitária (UFC), 1989.

HABERMAS, Jürgen. *Direito e democracia: entre facticidade e validade*, v. 1, São Paulo: Tempo Brasileiro, 1997.

_____. Fios do tempo: precisamos agir como o saber explícito do nosso não saber. *Ateliê de humanidades*, 12 abr. 2020; Disponível em: <ateliêdehumanidades.com/2020/04/12/fios-do-tempo-precisamos-agir-com-o-saber-explicito-de-nosso-nao-saber-entrevista-com-jurgen-habermas>. Acesso em: 12 set. 2020.

HAN, Byung-Chul. La emergencia viral y el mundo de mañana. *El País*, 21 mar. 2020; Disponível em: <elpais.com/ideas/2020-03-21/la-emergencia-viral-y-el-mundo-de-manana-byung-chul-han-el-filosofo-surcoreano-que-piensa-desde-berlin.html>. Acesso em: 12 set. 2020.

HUSSERL, Edmund. *A crise da humanidade europeia e a Filosofia*. Porto Alegre: Edipucrs, 2006.

MARCUSE, Herbert. *Cultura e sociedade*. v.1. São Paulo: Editora Paz e Terra, 1997.

_____. A Arte na sociedade unidimensional. In: LIMA, Luiz Costa (org.). *Teoria da cultura de massa*. São Paulo: Paz e Terra, 2005.

MBEMBE, Achille. The age of humanism is ending. *Mail & Guardian*, 22 dez. 2016. Disponível em: <mg.co.za/article/2016-12-22-00-the-age-of-humanism-is-ending>. Acesso em: 12 set. 2020.

_____. *Políticas da inimidade*. Lisboa: Antígona Editores, 2017.

_____. *Direito universal à respiração*. São Paulo: N-1, 2020.

SILVA, Virgílio Afonso da. O proporcional e o razoável. *Revista dos Tribunais*, São Paulo, v. 798, p. 23-50, 2002.

Um debate sobre o campo online e a etnografia virtual

Samira Cristina Silva Pereira¹

Sérgio Procópio Carmona Mendes²

Resumo: O objetivo deste artigo é apresentar a etnografia virtual como uma opção válida de pesquisa, sem desprezar a etnografia nos seus moldes clássicos. A sociedade conectada por uma rede mundial de computadores passa por uma reconfiguração. Assim, busca-se expor, a partir de uma revisão bibliográfica, tais transformações. A etnografia virtual entra em cena com suas inovações metodológicas, sendo, dessa forma, moldada para as necessidades do novo campo *online*. Ao estabelecer novas relações entre o pesquisador e os interlocutores e entre o pesquisador e o campo de pesquisa escolhido, este estudo faz-se necessário devido às lacunas dentro da literatura acerca da etnografia virtual, pouco explorada e que ainda enfrenta considerável resistência às inovações dos métodos da etnografia. Dessa forma, é indispensável compreender e conhecer melhor o método da etnografia virtual, pois, uma vez que vivemos imersos na tecnologia de comunicação mediada por computadores, os artefatos digitais ganham cada vez mais espaço na produção cultural e nas interações sociais, que passam a ocorrer pela mediação de máquinas em formatos digitais.

Palavras-chave: Tecnologia. Etnografia. Etnografia virtual. Antropologia. Método.

1 Graduada em Ciências Sociais pela Universidade Federal de Alfenas (UNIFAL-MG). Atualmente mestranda em Ciências Sociais pela Universidade Federal de Juiz de Fora (UFJF). cv Lattes: lattes.cnpq.br/3450683107337235. E-mail: samiracsp@outlook.com.

2 Doutor em Ciências Sociais pela Universidade Estadual de Campinas (UNICAMP); Mestre em Antropologia Social pela Universidade Estadual de Campinas (UNICAMP); Graduado em Ciências Econômicas pela Universidade Estadual Paulista (UNESP). cv Lattes: lattes.cnpq.br/4662751365319253. E-mail: serpcm@gmail.com.

A debate about the online field and virtual ethnography

Abstract: The aim of this article is to present virtual ethnography as a valid option of research without neglecting the methods of ethnography along their classical lines. Connected in a worldwide computer network, society is currently undergoing a deep reconfiguration. The paper offers a bibliographical outline of the ongoing transformations in the domain of virtual ethnography and its methodological innovations. In establishing new relationships between scholars and their interlocutors, between researchers and their research fields, this study aims at closing some of the gaps in the literature concerning virtual ethnography, a field of research so far little explored and still facing considerable resistance to methodological innovation. The paper also aims at contributing to a better understanding of the method of virtual ethnography in a world immersed in computer-mediated communication technology, with digital artifacts gaining more and more space in cultural production and social interaction conducted more and more through the mediation of machines in digital formats.

Keywords: Technology. Ethnography. Virtual ethnography. Anthropology. Method.

A maneira como percebemos o mundo vem se alterando devido às transformações proporcionadas pelos novos recursos tecnológicos. A partir de tais transformações, emerge a Sociedade da Informação, na qual estamos inseridos. Conforme Manuel Castells (1999), a tecnologia da informação se tornou uma ferramenta indispensável para a construção do conhecimento e também a transmissão de informação se torna a principal fonte de poder. O cenário social passa a ser alterado devido a que a tecnologia de informação passa a remodelar as bases da sociedade, impondo um ritmo acelerado de transformações. Abrem-se a partir daí portas para novas maneiras de se relacionar com o outro. Logo, abrem-se infinitas possibilidades de receber e gerar informação, que eram anteriormente inimagináveis. Para Paula Sabilia (2008), nossa sociedade está passando por uma época limítrofe, o que acarreta uma ruptura que atinge diversos âmbitos – políticos, socioculturais e econômicos – rumo a novos horizontes.

Tendo isso em vista, é possível destacar novas práticas, de acordo com a autora (ibid.). Elas ocorrem devido aos textos em formato eletrônico, lidos ou escritos, que correspondem a novas formas de diálogo digitado. Outras vezes, as interações se dão por meio de imagens fixas ou em movimento. Inauguram-se, assim, novos hábitos e novas vivências para leitores e para autores. A partir disso, despertou-se o interesse em compreender como tais mudanças impactam a metodologia de pesquisa no campo das ciências sociais, especificamente o método utilizado na Antropologia, a etnografia.

Convém destacar o posicionamento de Bernhard Rieder e Theo Röhle (2012), em seu trabalho *Digital Methods: Five Challenges*. Nele, os autores abordam o uso de computadores em pesquisas das ciências humanas, sendo possível ressaltar que os artefatos digitais ganham cada vez mais espaço na produção cultural e também nas interações sociais, que passam a ocorrer pela mediação de máquinas em formatos digitais. De acordo com os autores, é possível destacar como exemplo o surgimento de formulários digitais, *e-books*, softwares, entre outros, que alavancaram as pesquisas por meio de tais aparatos tecnológicos. Além disso, a cada dia a quantidade de material, informações e dados armazenados nas redes virtuais se amplia. É fundamental repensar a maneira de lidarmos com a análise desse volume de informações.

Diante dessa abundância de material empírico, os pesquisadores estão cada vez mais voltando-se para métodos automatizados de análise, a fim de explorar esses artefatos e as realidades humanas com as quais estão entrelaçadas. [...] As ferramentas são promissoras: são capazes de processar corpora muito maiores do que seria possível fazer manualmente; eles fornecem a capacidade de integrar perfeitamente zoom de micro para macro e sugerimos que podemos reconciliar a amplitude e profundidade de análise; eles podem ajudar a revelar padrões e estruturas que são impossíveis de discernir a olho nu. (RIEDER; RÖHLE, 2012, p. 67)

Introduzidas tais mudanças, o contexto e as transformações da sociedade contemporânea, o próximo passo será ressaltar o conceito de método. Esse conceito é destacado por Bernhard Rieder e Theo Röhle (2012) que, ao realizarem uma busca em um dicionário, conferem a ela o significado de método, que corresponde à busca pelo conhecimento ou modo de investigação. Dessa maneira, os autores afirmam que o método é uma busca ordenada e sistemática pelo conhecimento por meio de um processo de investigação. Ademais, o método está diretamente relacionado aos objetivos específicos do pesquisador, que acarretam decisões específicas e levam a descobertas conforme o problema proposto. Alguns exemplos de métodos nas ciências sociais são: observação participante, técnicas de regressão estatística, análise semiótica, entre outros.

Rieder e Röhle (2012) comparam o método com uma receita ou até mesmo com uma planta de um prédio ou casa. Tendo em vista que o método guia a pesquisa, a fim de compartilhar experiências na construção do conhecimento, o método também estabelece pontos de referência fornecendo um norte para o pesquisador, orientando-o nos caminhos por ele escolhidos. É neste momento que os aparatos e tecnologias digitais entram em cena como ferramentas ou campo para pesquisa.

Com o advento das novas mídias sociais, é necessário lançar um olhar antropológico sobre os impactos e efeitos que elas têm na relação dos sujeitos com seus aparatos digitais. Para tanto, faz-se necessário compreender o debate sobre o método, suas transformações nas últimas décadas e a distinção entre a etnografia clássica e o novo tipo de etnografia empregado para efetivar as análises destes novos fenômenos, ou seja, é preciso evidenciar a discussão sobre etnografia virtual, também chamada etnografia digital, e sobre como essa técnica de pesquisa é utilizada para coleta de dados no ciberespaço.

Em um primeiro momento convém olhar para a Antropologia clássica, ao buscar compreender o método etnográfico nos moldes de Malinowski (1884-1942), responsável por consagrar a técnica da observação participante. Essa técnica é citada por antropólogos ainda nos dias de hoje, que o consideram como referência, utilizando sua obra como manual

para pesquisa de campo. Em um segundo momento, foi feito um mapeamento de artigos que discutem a etnografia virtual, que é uma inovação dos métodos tradicionais da etnografia, devido às transformações citadas acima. O campo do antropólogo passa ser o ciberespaço, que é o cenário para interação dos grupos. Sendo assim, faz-se necessário compreender as adaptações que ocorreram no método clássico, e as possibilidades de se explorar a etnografia em um campo virtual.

Por fim, o intuito deste trabalho é ressaltar a importância de apreender o método da etnografia virtual, tendo em vista as mudanças que ocorreram na Sociedade da Informação e na urgência de se olhar para fenômenos no ambiente *online* e as reconfigurações sociais realizadas pelas redes e mídias sociais. As mídias sociais se referem à postagem de arquivos e informações do usuário, sem gerar relacionamento direto, e priorizam a velocidade e o dinamismo das interações, com diversos conteúdos; já as redes sociais correspondem a espaços virtuais que constroem laços estruturados a partir das conexões virtuais proporcionadas pela Internet (GASQUE, 2016). É importante compreender quais os processos de relacionamento dos sujeitos no ciberespaço, portanto, convém apontar as vantagens e desvantagens da etnografia virtual como ferramenta para investigar as novas relações que vêm sendo estabelecidas nas redes.

O método etnográfico

De início, faz-se necessário apresentar as principais características do método etnográfico. A etnografia consiste, portanto, em um conjunto de técnicas e práticas por meio das quais se opera a coleta de dados em um determinado campo, ou seja, o grupo social escolhido pelo pesquisador, o antropólogo.

A etapa introdutória corresponde à escolha do tema e do objeto pelo pesquisador, que necessita ter conhecimento prévio sobre a temática para recortar a problemática central a ser pesquisada. Desse modo, é preciso recorrer ao aporte teórico que será mobilizado, pois não se trata de um mero olhar para o outro, mas de diversas etapas construídas gradualmente a fim de apreender o fenômeno investigado.

Segundo Ana Luiza Rocha e Cornelia Eckert (2008), a etnografia corresponde a demandas científicas que ocasionam a produção de dados e conhecimentos antropológicos por meio da interação entre sujeito e pesquisador, assim propondo observação direta ou indireta por meio de conversas informais ou formais, que podem ocorrer também em entrevistas semiestruturadas ou não. Tais variações podem ocorrer devido aos interlocutores e ao campo escolhido pelo pesquisador.

Além da observação e das entrevistas, esse método é composto de diversas etapas que exigem um esforço do pesquisador de se abster previamente de juízos de valor, para tentar compreender as práticas e ritos de uma cultura específica sem julgamentos. De acordo com Magnani (2009), o primeiro desses esforços corresponde ao estranhamento ou exterioridade do pesquisador para com o objeto. O autor ainda destaca que, nessa etapa da pesquisa, o pesquisador adentra o universo dos interlocutores ao compreender sua visão de mundo por meio de uma relação de troca estabelecida em campo.

Magnani, além disso, observa que o pesquisador deve procurar pistas para formular suas teorias ou compará-las. Dessa maneira, por meio de “*insights*” que produzem o conhecimento, com um trabalho paciente e contínuo no qual a observação perspicaz resulta em novas percepções. Convém diferenciar “experiência etnográfica” de “prática etnográfica”: segundo Magnani (2009), a primeira corresponde a algo programado e contínuo, e a segunda corresponde à descontinuidade e é imprevista.

Posterior à coleta de dados em campo, o pesquisador começa a etapa da transcrição em forma de monografia, que consiste em estruturar e compartilhar o conhecimento acumulado durante os atos interativos com o outro. Essa construção se dá pela estruturação dos dados observados, da revisão bibliográfica de pesquisas sobre a temática, do grupo selecionado, das análises de discursos e das entrevistas. Em alguns casos, documentos históricos ou estudo de imagens, entre outros recursos, podem ser utilizados para a produção de dados empíricos e coletados. Rocha e Eckert (2008) observam que o ato de escrita etnográfica corresponde a “recriar” tais formas culturais que foram observadas, por meio de uma estrutura textual narrativa.

Compreendido o que é o método e os passos que o pesquisador deve seguir, cabe, portanto, discutir as premissas dos clássicos da Antropologia, visto que o fazer etnográfico é pautado na tradição. Dentre os fundadores, destaca-se o antropólogo Bronislaw Malinowski, considerado um dos pilares da disciplina de Antropologia. Ao inaugurar os fundamentos do método etnográfico, que viriam a servir de aporte para posteriores trabalhos antropológicos, Malinowski é considerado o fundador da Antropologia Social e da Escola Funcionalista. Segundo Roberto Cardoso de Oliveira (2014), a escola Funcionalista é uma vertente de pesquisa em antropologia que propõe explicar os fenômenos sociais por meio de suas funções, nas quais cada elemento cultural poderia ser destinado a cumprir determinada função dentro da estrutura social. Como fundador dessa

escola, Malinowski, em sua obra “Os Argonautas do Pacífico Ocidental” (1978), explica como foi desenvolvido seu trabalho de campo, fornecendo detalhes sobre a construção metodológica. Para ele o trabalho etnográfico obtém valor científico quando os resultados podem ser obtidos por meio de observação direta.

Ao respaldar este estudo nas perspectivas de Malinowski, o intuito é compreender a etnografia nos moldes clássicos. Segundo ele, pode-se obter sucesso na aplicação sistemática de regras e princípios científicos:

Os princípios metodológicos podem ser agrupados em três unidades: em primeiro lugar, é lógico, o pesquisador deve possuir objetivos genuinamente científicos e conhecer os valores e critérios da etnografia moderna. Em segundo lugar, deve o pesquisador assegurar boas condições de trabalho, o que significa, basicamente, viver mesmo entre os nativos, sem depender de outros brancos. Finalmente, deve ele aplicar certos métodos especiais de coleta, manipulação e registro da evidência. (MALINOWSKI, 1978, p. 20)

Compreende-se, a partir da leitura da obra de Malinowski (1978), que o objetivo da pesquisa etnográfica em campo consistia em compreender a constituição tribal e os fenômenos culturais, a fim de delinear padrões e leis gerais. Dessa maneira, a pesquisa etnográfica extrai do fenômeno cultural observado regularidades relevantes, e elas podem ser captadas pelo pesquisador a fim de produzir material científico.

Eunice Durham (1978) analisa a obra de Malinowski, destacando os movimentos realizados pelo autor ao construir a etnografia pela vivência em campo. Para ela, a vivência fornece materiais para o pesquisador por meio das situações de pesquisa em meio à sociedade sob estudo. No decorrer da investigação, o pesquisador coleta material ao lançar sua visão particular sobre o objeto de pesquisa utilizando o método.

Malinowski proporcionou, assim, a articulação do objeto e método etnográfico. Durham (1978) expõe ainda que o marco de Malinowski consiste na descoberta da observação participante. Essa técnica propõe a interação entre observador e observado. Assim, é possível acompanhar o interlocutor ao apreender e observar seus costumes e crenças. A observação participante, portanto, consiste em participar da vida social do observado.

A partir dos trabalhos de Malinowski, instaurou-se a observação participante como uma exigência para o trabalho de campo em Antropologia. Conforme Durham (1978), é um processo de transformação do observador ao assimilar categorias do universo cultural estudado que necessitam ser ordenadas. Portanto, o próprio observador realiza um esforço ao tornar-se instrumento de sua pesquisa de campo por meio de

processos sistemáticos, como, por exemplo, aprender a língua falada por aqueles interlocutores. Essa transformação do pesquisador pode ser uma das maiores dificuldades enfrentadas, devido à necessidade de distanciar o lado pessoal de sua produção científica em campo.

Mariza Peirano (1995) aborda a discussão acerca da etnografia de maneira distinta, ao afirmar que, apesar de possuir tradições sólidas, a Antropologia dispõe de distintos estilos de pesquisa, que podem influenciar os resultados obtidos, bem como a personalidade do pesquisador. Sendo assim, “a experiência de campo depende, entre outras coisas, da biografia do pesquisador, das opções teóricas dentro da disciplina, do contexto sócio-histórico mais amplo e, não menos, das imprevisíveis situações que se configuram, no dia-a-dia” (PEIRANO, 1995, p. 22).

É necessário, portanto, um equilíbrio entre teoria e prática na pesquisa antropológica. Desse modo, não há prática sem fundamentação teórica e não há teoria sem o trabalho de campo, no qual a teoria é colocada à prova, comparando e buscando responder a questões formuladas pela teoria. Para além disso, Peirano (1995) afirma que a prática é essencial nas posições que o autor escolher em campo.

Uma boa etnografia, ainda de acordo com Peirano (1995), necessita ser rica em dados iniciais suficientes que possibilitem uma reanálise, descrevendo o método como sendo artesanal, microscópico e minucioso. Portanto, os dados coletados e as informações obtidas devem estar para além de afirmar um ponto de vista teórico, mas para indicar que “haverá sempre a ocorrência de novos indícios, dados que falarão mais que o autor e que permitirão uma abordagem diversa” (PEIRANO, 1995, p. 52). Assim, a obra de um antropólogo não é linear, mas construída em determinado contexto a partir de peculiaridades biográficas, como a obra de Malinowski, em que o bom texto etnográfico é um experimento.

Finalmente, a próxima etapa a ser discutida é a etnografia virtual. Uma forma recente de análise de dados etnográficos, que explora o campo do ciberespaço. Inaugura, dessa forma, um tipo de prática e análise etnográficos ainda pouco explorados pelos pesquisadores da Antropologia.

Etnografia Virtual

Devido às transformações expostas, começam a surgir debates sobre as formas de produção do conhecimento que dialoguem com relações sociais do mundo contemporâneo (GUSMÃO, 2008), visto que a Antropologia tradicional, em seus primórdios, consagrou-se na classificação de

grupos considerados primitivos ou atrasados, tribais ou pré-modernos. A autora enfatiza, ainda, ser esse um campo de tensão dentro da disciplina. Por um lado, as trajetórias do antropólogo devem averiguar melhor as diferenças sociais, éticas e propor novas formas de compreender a realidade social. Por outro, torna-se insuficiente a tradição para compreender o contexto político das diferenças.

Tornou-se, portanto, traço distintivo da Antropologia seu caráter classificatório e também descritivo, já que essas eram características necessárias para se observar os povos em colonização nos séculos XIX e XX. Gusmão (2008, p. 50) aponta que, ao constituir-se, “o fazer antropológico, nessa medida, era o de submeter e colocar em submissão o outro e seu mundo, tendo por meta a civilização e a humanidade”. A Antropologia, como ciência, era outrora tida como descrição ou transformação do “mundo do outro”.

Mas o que precisa ser enfatizado atualmente é o jogo entre ciência e prática, que são alvo de constantes discussões. Ao constatar as transformações que ocorreram nas últimas décadas, é possível destacar também as transformações dentro do mercado de consumo, no qual a etnografia pode ser útil no intuito de compreender melhor a realidade dos sujeitos e do seu cotidiano. Esse método desperta interesse de profissionais da área do marketing, no que diz respeito à consultoria e pesquisa de mercado. Assim, a disciplina passa a ser vista como uma ferramenta para compreender mudanças sociais, políticas e culturais, conforme Magnani (2009). O autor discorre, ainda, como exemplo, acerca dos movimentos sociais urbanos, que tiveram início por volta de 1968. Devido à repressão do golpe militar, mudanças no cenário político e na habitação das cidades puderam ser notadas e, entre elas, a exigência por melhores condições de vida nesse novo cenário.

Posteriormente a essas mudanças que ocorreram em meados da década de 1970, os antropólogos passaram a ter problemas de ordem teórico-metodológicas, conforme Magnani (2009), visto que anteriormente as fronteiras por eles pesquisadas eram bem delimitadas. A Antropologia, então, passou a voltar seus olhares para atores sociais e suas práticas no contexto sociopolítico. No entanto, as novas utilizações da disciplina podem trazer uma banalização metodológica, que corresponde à utilização indevida do método etnográfico, por meio de uma postura intelectual inadequada ao utilizar as estratégias de pesquisa (MAGNANI, 2009).

Cabe à etnografia um novo olhar, de acordo com Magnani (2009), “*de perto e de dentro*” (grifo do autor), ou seja, olhar para os arranjos dos atores sociais (LATOUR, 2012). Esses atores correspondem ao poder público, corporações privadas, moradores, visitantes, equipamentos, entre

outros. Por meio das redes de interações, ocorrem trocas e conflitos. Conforme o autor, “esse resultado, sempre em processo, constitui, por sua vez, um repertório de possibilidades que, ou compõem o leque para novos arranjos ou, ao contrário, surgem como obstáculos. Cabe à etnografia captar esse duplo movimento” (MAGNANI, 2009, p. 133).

A tecnologia e seus aparatos, juntamente com a Antropologia, podem contribuir para compreender relações de troca, conflitos e interações que na última década passaram a ocorrer no meio virtual. Portanto, faz-se necessária uma abordagem para explorar a Internet e o ciberespaço como campo de pesquisa, a fim de entender nossas atitudes perante as mudanças tecnológicas, tais como aprendizagem, lazer, entre outras.

Portanto, a etnografia virtual corresponde a um método de analisar dados perante as mudanças tecnológicas, sendo indispensável tratar da distinção dos termos empregados para designar esse campo, que pode ter como sinônimo: “netnografia”, “etnografia digital”, “webnografia” e também “ciberantropologia”. Christine Hine (2000) é a principal responsável pela popularização do termo “etnografia virtual”. Outra variação de terminologia corresponde a “netnografia”, utilizado inicialmente durante os anos 90, sendo relacionado ao marketing e às comunidades de consumo online. Além disso, há outros termos utilizados que surgiram do flerte entre pesquisas acadêmicas e de mercado que, conforme Frago, Recuero e Amaral (2011), seriam meramente definições operatórias, tais como “etnografia digital” e “webnografia”. Segundo os autores, esses são neologismos que podem ser incorporados à ciberantropologia.

Hine (2000) é precursora nos estudos etnográficos sobre Internet, e busca compreender o *status* da Internet e o contexto do seu uso. Em “*Virtual Ethnography*”, publicado em 2000, a autora compreende a Internet por meio de dois vieses. No primeiro, ela é entendida como um lugar, ou ciberespaço, no qual a cultura é constituída. E no segundo, como artefato cultural, a Internet como um produto da cultura com objetivos e propriedades distintos conforme o contexto em que está inserido.

Em defesa do método da etnografia virtual, Hine afirma que ele pode ser eficiente para compreender os meios tecnológicos e as culturas concebidas pela Internet. E, por meio desse método, compreender também as complexas relações estabelecidas pelas tecnologias em diferentes esferas, além da Internet. Sendo assim, a etnografia realizada na Internet visa atentar para detalhes, de acordo com Hine (2000). Dessa maneira, a autora enfatiza os usos e apropriações que os atores sociais realizam por meio da Internet, e que possuem diversos significados culturais. Não obstante, o surgimento de outra perspectiva destaca a Internet com o importante o papel na construção de práticas sociais e performances (POLIVANOV, 2013).

Ao apontar características do método e os graus de inserção do pesquisador durante a pesquisa, percebe-se que essa pode ocorrer de duas maneiras: o *lurker* e o *insider* (FRAGOSO; RECUERO; AMARAL, 2011). O primeiro corresponde àquele pesquisador que se limita somente às observações do grupo, interferindo o mínimo possível nas práticas observadas, uma prática denominada *lurking*, que significa ‘ficar à espreita’. Tal prática seria característica do ciberespaço e, por meio dela, o ator não se manifesta, apenas dedicando-se à observação do comportamento dos outros (POLIVANOV, 2013, p. 64). Em oposição, está o pesquisador *insider*, técnica que diz respeito à proximidade do pesquisador com seus interlocutores.

Os principais problemas metodológicos da inserção do pesquisador correspondem ao informante conduzir a pesquisa de maneira confusa ou não relatar experiências reais, tendo em vista que as entrevistas e relatos precisam ser feitos com bastante atenção por parte do pesquisador, que deve manter um olhar crítico e afastado do objeto, com rigor ao sistematizar os dados coletados e descrições. Isso demanda atenção redobrada, pois publicar determinados relatos e experiências pode afetar a relação do pesquisador e informante. Além disso, o pesquisador *lurker* precisa entrar no grupo de alguma forma, sendo percebida sua presença em alguns momentos pelos interlocutores.

Por meio dessas duas maneiras de inserção do pesquisador, a etnografia virtual pode ser conduzida a partir de duas perspectivas. O distanciamento, realizado pelo *lurker*, que faz a observação das interações dos atores sociais no ciberespaço a partir de uma “observação não participante” – termo utilizado por Polivanov (2013) – em determinado ambiente virtual, em que o pesquisador coleta dados como textos, também imagens e *emoticons*, sem realizar interferência no meio. E o envolvimento, realizado pelo *insider*, em que a observação é participante e o pesquisador interage com seu interlocutor.

Essa última, de acordo com Polivanov (2013), afirma a mediação entre o pesquisador e o interlocutor. Mas dentro do ambiente virtual, essa relação tem características próprias, como por exemplo a gramática e a linguagem utilizadas em rede. Ao observar uma discussão ou um grupo na Internet, seja em um site ou rede social, o pesquisador deve estar atento para textos escritos, *emoticons*, imagens ou até mesmo *links* publicados, entre outros. Assim, a atenção do pesquisador da etnografia *online* se orienta para detalhes distintos daquela realizada de forma presencial.

A etnografia virtual, portanto, busca compreender o interlocutor do ciberespaço. Esse é o campo que promove a pesquisa em um meio de comunicação composto por máquinas, cenários, atores e programas. Seu modo de funcionamento se dá a partir de interpretação de dados, infor-

mações, reproduzidas por computadores e redes. De acordo com Lévy (1999), ocorre por meio da interconexão mundial de computadores, a Internet, que afirma que o ciberespaço corresponde a um dispositivo de comunicação interativa, à sociabilidade, organização, transação, por meio de um mercado da informação e conhecimento.

Dentro desse universo, a cibercultura exerce papel fundamental, já que propõe perspectivas inimagináveis anteriormente. Lévy (1999) afirma que corresponde, portanto, a um conjunto de técnicas de comunicação que se configuraria nessa forma de se relacionar com o outro. Assim, a interconexão de mensagens que ultrapassam distâncias configuram formas diversas de vinculação e circulação dos conteúdos. O ciberespaço e a cibercultura, portanto, são resultantes dessa reconfiguração social, e abrem portas para novos estilos de relacionamento.

Em suma, a partir de tais mudanças orquestradas na sociedade contemporânea pela cibercultura, convém aos etnógrafos adequar as técnicas de pesquisa para as distintas abordagens a serem realizadas, em que os mundos *online* e *offline* se fundem cada vez mais, nas diversas esferas da sociedade. Ambos os mundos colidem, sendo transformados por meio das interações. Faz-se necessário pensar tais perspectivas de mudanças por meio do olhar antropológico, ao surgir “a necessidade de repensar a ideia de agência e hibridismo, superando a dicotomia humano/não humano, especialmente no campo da chamada ‘comunicação mediada por computador’” (RIFIOTIS, 2016, p. 88).

É indispensável incorporar ao debate as perspectivas de Donna Haraway (2000); para ela, somos todos híbridos, teóricos e fabricados por meio de máquinas e organismos; segundo ela, somos “ciborgues”, isto é, um organismo cibernético, híbrido de máquina e ser biológico, que corresponde a uma criatura da realidade social e da ficção. A técnica utilizada pela etnografia virtual permite analisar esse contexto social a fim de compreender a conexão entre humanos, mediada por não humanos.

Não obstante, pesquisadores ortodoxos, defensores de métodos tradicionais *offline*, questionam os métodos propostos pela etnografia virtual. Questionam também o ciberespaço como lugar no qual o pesquisador e o pesquisado podem ter sua relação intermediada por máquinas. Para esses pesquisadores, uma etnografia “de fato” não poderia se dar no ciberespaço, uma vez que, para eles, faltam “o deslocamento, o estranhamento e o ‘ir a campo’ tão decisivos na formação do olhar interpretativo” (POLIVANOV, 2013, p. 65-66). Em oposição, autores já citados acima, inclusive Hine (2000), defendem que a etnografia realizada em um ambiente virtual é rica em significados:

Uma vez que pensemos o ciberespaço como um lugar onde as pessoas fazem coisas, nós podemos começar a estudar exatamente o que é que elas fazem e porque, nos seus termos, elas o fazem. No entanto, assim como com todas as metodologias, mover a etnografia para um ambiente *online* tem envolvido algumas reexaminações do que a metodologia implica. (HINE, 2000, p. 21)

A etnografia virtual é importante para resolver problemas de pesquisa que estão relacionados à Internet, vista “como um espaço social no qual se pode legitimamente fazer pesquisa antropológica” (RIFITIS, 2016, p. 88), uma vez que as mídias passaram a protagonizar uma articulação de práticas, crenças, rituais e modos de estar no mundo.

No intuito de facilitar a compreensão acerca desse tipo de etnografia, a melhor forma é exemplificar como funciona o método com uma pesquisa já realizada. Bonilla e Rosa (2015) analisaram, no *Twitter*, a presença e expressão de uma *hashtag*. A *hashtag* que os autores optaram por pesquisar foi *#Ferguson*, devido ao tiroteio que matou o jovem negro nos EUA em 10 de Agosto de 2014, acontecimento divulgado por meio do *Twitter* e do jornalismo informal, gerando um grande protesto na rede social. Essa plataforma é destacada pelos autores como meio de engajamento político, na qual eles examinaram as formas de ativismo em rede. Assim, eles inauguram uma forma de fazer etnografia no meio virtual, inovando seus métodos de acordo com a temática de interesse, adaptando o método para redes sociais utilizadas nos dias de hoje.

Bonilla e Rosa (2015) afirmam que o *Twitter* corresponde a um “não lugar”, local transitório correspondente a uma instância do ciberespaço. Dessa maneira, os autores questionam: esse tipo de Antropologia seria fazer uma “Antropologia de poltrona” ou seria um passo em direção à Antropologia do século XXI, cada vez mais preocupada com as práticas digitais e seus fenômenos? Os autores constatam ainda que os antropólogos interessados nesses mundos sociais devem permanecer atentos para as possibilidades de pesquisa utilizando a “*hashtag* Etnografia”, preparando-se para ler e ir além das linhas digitais, visto que a *hashtag*, segundo eles, é uma janela que oferece a oportunidade para entender melhor o virtual, através dos *tweets* dos membros da rede social *Twitter* e seus verdadeiros engajamentos.

Os computadores, portanto, são utilizados para uma grande variedade de pesquisas. Pode-se citar o exemplo da Universidade de Londres, de acordo com Rieder e Röhle (2012), sobre as técnicas de análise textual empregadas para classificar suas estruturas. Os autores assinalam ainda como tais pesquisas facilitam trabalhar com grandes coleções multimídia:

A pesquisa baseada em software nos espaços digitais “nativos” da Web se expandiu a um ritmo impressionante e o grande volume de dados processados tornou-se bastante assustador. Embora o bem citado estudo de Adamic e Glance (2005) sobre a blogosfera política nos Estados Unidos em 2004 colheu continuamente 1.000 blogs, pesquisadores do *Orange Labs* na França (Prieur *et al.* 2008) examinaram todas as cinco milhões de contas de usuário público no *Flickr* em 2006 e um dos muitos estudos recentes no *Twitter* (Cha *et al.* 2010) analisaram as interações entre 54 milhões de perfis no serviço de *microblog*. Enquanto alguns desses projetos de pesquisa parecem mais intimamente aliados à física do que às humanidades em mentalidade e metodologia, é óbvio que eles representam novas possibilidades estudar a interação e a imaginação humanas em uma escala muito grande. (RIEDER; RÖHLE, 2012, p. 68)

As mídias digitais fornecem um oceano de dados para o etnógrafo realizar o trabalho de campo. Portanto, ao realizar a observação participante no meio virtual, a natureza da observação muda. E também as formas de construir um caderno de campo passam a ser mais tecnológicas, registrando eventos, interações e locais digitalmente. Assim, as observações participantes virtuais incluem interações em salas de bate papo, *e-mails*, mensagens, *sites*, entre outros. Contudo, esse tipo de observação consiste em observar textos e imagens por meio da tela do computador por meio da participação ativa em determinado grupo pesquisado, e na oportunidade de observar os grupos *online*, além de postar e participar nos mesmos.

Santos e Gomes (2013), em sua pesquisa, verificaram como os estudiosos da cibercultura estão se apropriando do método. Eles mapearam teses e dissertações do Banco de Teses e Dissertações do Instituto Brasileiro de Informação Científica e Tecnológica – BDTD/IBICT. Dentre as obras, 35 trabalhos foram selecionados e foi feita novamente uma seleção na qual 7 foram selecionados para a análise da metodologia. Questões como, por exemplo, neologismos em relação a terminologias tradicionais foram destacadas, e apontam a necessidade de se adotar uma nomenclatura demarcada. Uma fragilidade da etnografia digital foi a falta de rigor com termos e consentimentos.

Os autores assinalam ainda que, em grande parte das pesquisas, o método está sendo aplicado às redes sociais, o que se infere por meio da expansão dessas mídias, devido à praticidade e ao seu grande número de integrantes. Os autores ressaltam que o tempo de imersão e participação do pesquisador influencia na validação da pesquisa. Outro ponto importante destacado pelos autores corresponde ao posicionamento que o pesquisador deve ter ao explicitar a forma de sua entrada em campo, se a observação é participante e se há interação com os grupos pesquisados.

Por fim, relatam também que as entrevistas são ferramentas bastante utilizadas, em grande parte semiestruturadas. No intuito de esclarecer questões específicas por meio das interações, certos cuidados e preocupações devem ser apresentados no fazer etnográfico, para não violar a intimidade e a privacidade dos interlocutores.

Conclusão

É possível concluir que as tecnologias digitais são portas para um universo na pesquisa antropológica. Para além de meros processadores de texto, os computadores e as informações armazenadas nas redes são dados para as pesquisas e, além disso, correspondem a um fenômeno social que emergiu na Era da Informação.

A etnografia virtual é, portanto, um método, uma ferramenta que pode ser muito útil para o antropólogo ao analisar textos *online*, imagens, sons entre outros dados e fenômenos do ciberespaço. Mas para tal, o método deve ser utilizado pelo pesquisador com certos cuidados, pois ele precisa adaptar-se a essa ferramenta, realizando uma análise mais crítica.

Assim, transformações são necessárias para que o pesquisador consiga compreender e analisar o que ocorre dentro desse universo digital e como se dão as relações mediadas por computador. Apesar de tal adaptação ser fundamental para a disciplina de Antropologia, foi possível notar que são diversos os obstáculos para a adesão ao método da etnografia virtual, visto que aqueles pesquisadores que defendem os métodos tradicionais assinalam as dificuldades em adentrar ao campo, e questionam também se esse seria realmente um campo a ser explorado pelo fazer etnográfico. No entanto, notou-se que, apesar de se distinguir do método clássico, a etnografia virtual propõe diversas maneiras de se estabelecer contatos com seus interlocutores na forma de *insiders* ou *lurkers*. Cabe ao pesquisador adaptar-se a partir do seu objeto e problemática.

O pesquisador precisa ainda atentar à ética, havendo uma preocupação com a intimidade e privacidade dos sujeitos, que precisam estar sempre em primeiro lugar em uma pesquisa de campo que envolve interlocutores. Nesse sentido, propõe-se encorajar a exploração aos pesquisadores que se interessam por pesquisas no campo digital das ciências humanas e sociais. Vale a pena enfrentar os desafios e as armadilhas metodológicas nesse campo, para que nasçam pesquisas que venham a contribuir para a proliferação e construção do conhecimento científico.

Referências

BONILLA, Yarimar; ROSA, Jonathan. #Ferguson: Digital protest, hashtag ethnography, and the racial politics of social media in the United States. *American Ethnologist*, Hoboken, NJ, v. 42, n. 1, p. 4-16, 2015.

CASTELLS, Manuel. *A sociedade em rede*. Trad. Roneide Venancio Majer. São Paulo: Paz e Terra, 1999.

COLEMAN, Gabriela. Ethnographic approaches to digital media. *Annual Review of Anthropology*, Palo Alto, CA, v. 29, p. 487-505, 2010.

DURHAM, Eunice Ribeiro. *A reconstituição da realidade*. São Paulo: Ática, 1978.

FARIAS, Lídia; MONTEIRO, Taís. A identidade adquirida nas redes sociais através do conceito de persona. In: *XIX Prêmio Expocom*, Chapecó, 2012. Disponível em: <intercom.org.br/papers/regionais/nordeste2012/resumos/R32-1497-1.pdf> Acesso em: 24 nov. 2018.

FRAGOSO, Suely; RECUERO, Raquel; AMARAL, Adriana. Abordagens etnográficas. In: FRAGOSO, Suely; RECUERO, Raquel; AMARAL, Adriana. *Métodos de Pesquisa para Internet*. Porto Alegre: Sulina, 2011, p. 167-204.

GARCIA, Angela; STANDLEE, Alecea; BECHKOFF, Jenifer; CUI, Yan. Ethnographic approaches to the Internet and computer-mediated communication. *Journal of Contemporary Ethnography*, Thousand Oaks, CA, v. 38, n. 1, p. 52-84, 2008.

GASQUE, Kelley Cristine Gonçalves Dias. Internet, mídias sociais e as unidades de informação: Foco no ensino-aprendizagem. *Brazilian Journal of Information Studies: Research Trends*, Brasília, v. 10, n. 2, p. 14-20, 2016.

GUSMÃO, Neusa Maria Mendes. Antropologia, estudos culturais e educação: desafios da modernidade. *Pro-Posições*, Campinas, v. 19, n. 3, 2008.

HARAWAY, Donna. Manifesto Ciborgue: ciência, tecnologia e feminismo-socialista no final do século XX, trad. Tomaz Tadeu. In: HARAWAY, Donna; KUNZRU, Hari; TADEU, Tomaz (org.). *Antropologia do ciborgue: as vertigens do pós-humano*. Belo Horizonte: Autêntica, p. 33-118, 2000.

HINE, Christine. *Virtual Ethnography*. London: SAGE, 2000.

LATOUR, Bruno. *Reagregando o social: uma introdução à Teoria do Ator-Rede*. Salvador/Bauru: EDUFBA/EDUSC, 2012.

LEVY, Pierre. *Cibercultura*. São Paulo: Ed. 34, 1999.

MAGNANI, José Guilherme Cantor. Etnografia como prática e experiência. *Horizontes Antropológicos*, Porto Alegre, v. 15, n. 32, p. 129-156, 2009.

MALINOWSKI, Bronislaw. *Argonautas do Pacífico Ocidental*. São Paulo: Ática, 1978.

OLIVEIRA, Roberto Cardoso. O mal-estar da ética na Antropologia Prática. In: CERES, Vítora; OLIVEN, Ruben George; MACIEL, Maria Eunice; ORO, Ari Pedro (org.) *Antropologia e Ética: O debate atual no Brasil*. Niterói: Associação Brasileira de Antropologia e Editora da Universidade Federal Fluminense, 2014. p. 33-45.

PEIRANO, Mariza. *A favor da etnografia*. Rio de Janeiro: Relume-Dumará, 1995.

POLIVANOV, Beatriz Brandão. Etnografia virtual, netnografia ou apenas etnografia? Implicações dos conceitos. *Esféras*, Brasília, v. 2, n. 3, p. 61-71, 2013.

RIEDER, Bernhard; RÖHLE, Theo. Digital methods: Five challenges. In: BERRY, David (ed.) *Understanding digital humanities*. Houndmills: Palgrave Macmillan, 2012. p. 67-84.

RIFIOTIS, Theophilos. Etnografia no ciberespaço como “repovoamento” e explicação. *Revista Brasileira de Ciências Sociais*, Florianópolis, v. 31, n. 90, p. 85-98, 2016.

ROCHA, Ana Luiza Carvalho da; ECKERT, Cornelia. Etnografia: Saberes e práticas. *Iluminuras*, Porto Alegre, v. 9, n. 21, 2008.

SABILIA, Paula. *O show do Eu: intimidade como espetáculo*. Rio de Janeiro: Nova Fronteira, 2008.

SANTOS, Flávia; GOMES, Suely. Etnografia virtual na prática: análise dos procedimentos metodológicos observados em estudos empíricos em cibercultura. In: *Associação Brasileira de Pesquisadores em Cibercultura. Anais do VII Simpósio Nacional da Associação Brasileira de Pesquisadores em Cibercultura*. Disponível em: <abciber.org.br/simposio2013/anais/pdf/Eixo_I_Educacao_e_Processos_de_Aprendizagem_e_Cognicao/26054arq02297746105.pdf>. Acesso em: 13 abr. 2018.



RESENHAS

MARIN, Ronaldo; AMARAL, Gustavo Rick. *Os avanços da ciência ponderiam acabar com a filosofia?* São Paulo: Estação das Letras e Cores, 2020.

dx.doi.org/
10.23925/1984-3585.2020i21p214-217

Resenha do livro *Os avanços da ciência poderiam acabar com a filosofia?*, de Ronaldo Marin e Gustavo Rick Amaral

Peter A. Schulz¹

Ciência e filosofia caminham juntas. Essa é, de forma simplificada, a resposta à pergunta que intitula o livro resenhado: *Os avanços da ciência poderiam acabar com a filosofia?* Seus autores, Ronaldo Marin e Gustavo Rick Amaral, provocam a questão, que é destrinchada ao longo do texto com uma necessária, oportuna e, agora, urgente, densidade e fluência, tanto para o público mais especializado, quanto, principalmente, para o público geral.

O livro é necessário, porque aborda um problema central em como a ciência pode ser apresentada.

Há dois modos básicos para se explicar como a ciência chegou ao ponto em que estamos. Uma maneira é bem fácil de ser entendida, bem simples para ser explicada, mas provavelmente errada. A outra, é mais difícil de ser entendida, mais complicada para ser explicada, mas é provavelmente a mais correta. (MARIN; AMARAL, 2020, p. 20)

A maneira simples é a que normalmente é difundida, descrevendo a ciência como progresso cumulativo, contínuo e imutável. É a ciência apresentada através de seus resultados maravilhosos e sempre capaz de resolver os problemas que porventura surjam. A maneira mais complexa é a ciência descrita como seu processo, do difícil caminho de constante embate, de uma atividade humana que, na analogia dos autores, *troca a roda com o carro em movimento*. O cuidado de usar a palavra provavelmente para anunciar essa segunda maneira é devido à escolha utilizada para descrever, tanto a troca, quanto o movimento. Podem existir outras, mas embarcar nesse passeio vale a pena. Antes de comentar essa viagem, a oportunidade e a urgência merecem breves alusões.

¹ Doutor em Física, professor titular da faculdade de Ciências Aplicadas da Unicamp, professor e pesquisador em Estudos Quantitativos de Ciências.
cv Lattes: lattes.cnpq.br/6073974259027393. E-mail: peter.schulz@fca.unicamp.br.

O livro é oportuno pelo ambiente de crescente negação e ataque à ciência nos últimos tempos, antecipando que parte desse cenário surge no embate entre ciência e religião abordado no livro no terceiro capítulo. E a urgência vem da época de lançamento deste livro em meio à pandemia da Covid-19, durante a qual a valorização da ciência cresce, mas também novos ataques, ancorados a partir da negação da visão simplista da ciência e não da mais complexa, que é desenvolvida paulatinamente ao longo de quatro capítulos, que podem ser lidos (por que não?) como uma boa história de suspense.

Os autores do livro constroem a história da ciência utilizando-se de uma chave, que abarca também as causas de resistências à ciência em diferentes épocas: o conceito de unificação conceitual. A primeira unificação é a da mecânica clássica, quando, resumidamente, Isaac Newton unificou a descrição dos movimentos dos objetos cotidianos com os daqueles que compõem o resto do universo. A segunda unificação conceitual é a proporcionada pela Teoria da Evolução de Darwin, seres humanos e os outros animais são entidades da mesma classe: os primeiros não são especiais em relação aos segundos. A terceira unificação é a do corpo e mente, que está em curso, e que precisa ser comentada mais detalhadamente. A apresentação do problema e das duas primeiras unificações ocupam os dois primeiros capítulos. As questões colocadas pela segunda unificação, a Teoria da Evolução, é o gancho para o embate ciência e religião do terceiro capítulo, enquanto os desafios colocados pela terceira unificação conceitual são colocados no quarto e último.

O livro descreve o início da construção da ciência, quando a *troca da roda com o carro em movimento* criou o imaginário, ainda dominante, do mecanicismo do mundo: a descrição de todas as coisas pela ciência como engrenagens de um universo mecânico, mas com uma trava importante colocada pelo filósofo e cientista René Descartes: a alma seria impenetrável pela ciência. Essa trava perdurou por séculos e agora, com o desenvolvimento, por exemplo, das neurociências, a ciência começa a poder descrever, além das engrenagens do corpo, as engrenagens que Descartes descrevia como alma. Mas esse movimento de unificação conceitual rompe não apenas a trava, mas o próprio imaginário mecanicista, pois o que os autores observam como reducionismo, ou seja, a ideia de que o universo pode ser descrito a partir de um conjunto de regras/leis fundamentais, se opõe à complexidade: existem arcabouços de complexidade crescente nas diferentes *camadas de cebola* da ciência, da Física à Sociologia, passando pela Química, Biologia e todas as outras. Cada camada com seu

embate epistemológico, isto é, de como o conhecimento é construído e, mais contundentemente, como as camadas poderiam conversar entre si. Derrubar a trava de Descartes, requer enfrentar essas disputas. E aqui o diálogo entre ciência e filosofia não pode ser prescindido.

A leitura do livro evoca referências nele mencionadas, mas também outras, que de certa forma complementam a tese do livro. A proposta (equivocada) de obsolescência da filosofia, que é provocada pelo título e apresentada no início do livro é advogada, em boa parte, pelos mesmos que anteviam o *fim da ciência* com a elaboração da *teoria de tudo*, após a qual, ciência seria mera aplicação de um edifício de conhecimento sem novas grandes descobertas ou desafios (SCHULZ, 2018). É o que o físico Stephen Hawking (1942-2018), mencionado bem no começo do livro, defendia até alguns anos antes de falecer. Hawking também acabou se rendendo à emergência (os fenômenos que emergem em cada novo grau de complexidade e que não podem ser previstos pelas leis fundamentais da ciência no grau anterior. Como escreveu o físico Philip W. Anderson (1923-2020) em um manifesto publicado em 1973: “a habilidade de reduzir tudo a leis fundamentais simples não implica na habilidade de, a partir dessas leis, reconstruir o universo” (ANDERSON, 1973, p. 393).

Quanto ao embate ciência e religião, os autores declaram que não acreditam que seja possível uma conciliação entre essas formas de conhecimento, embora religiosos e cientistas possam perfeitamente se respeitar. Mas a dificuldade de conciliação fica clara no famoso debate entre Umberto Eco, pela ciência, e o cardeal Carlo Mantini, debate que ficou conhecido como *Em que creem os que não creem?* A dificuldade a mais para uma conciliação é comentada pelo professor de Estudos Bíblicos da Eastern University em St. Davids, Pennsylvania, Peter Enns: “Por que há tantos evangélicos em total alerta contra a evolução? Porque eles temem que, se a evolução é correta, suas heranças evangélicas sejam colocadas em xeque. Suas narrativas são ameaçadas” (ENNS, 2012). A grande questão colocada presentemente é que grupos anticiência começaram a atacar a narrativa da ciência para preservar as suas, como se Umberto Eco e Carlo Mantini não pudessem sequer conversar. Mais uma vez o livro de Ronaldo Marin e Gustavo Rick Amaral é bem-vindo.

Deixando o embate do parágrafo anterior de lado, resta perguntar: mas afinal, o restabelecimento de um diálogo entre ciência e filosofia só se apresenta como necessário na terceira unificação conceitual? A resposta, endossando a tese do livro, é não. Outros movimentos de tentativas de convergência vêm despontando. Um exemplo é o recente seminário

conjunto entre físicos e filósofos para resolver o beco sem saída em que uma área até então promissora da Física, a Teoria das Cordas (uma tentativa de unificação conceitual entre os dois pilares da Física moderna: a Teoria da Relatividade Geral e a Mecânica Quântica), em que dificuldades de progressos científicos se aliam a graves problemas epistemológicos e na qual os cientistas se renderam à filosofia. Como comentou em um seminário recente o prêmio Nobel de Física David Gross: “físicos precisam de filósofos e historiadores da ciência assim como pássaros precisam de ornitólogos.” (WOLCHOVER, 2015). No entanto, não estou mais certo de que não precisamos uns dos outros agora.

Os comentários dos três últimos parágrafos foram selecionados de uma série de ideias e associações que a leitura do livro provocou. Recomendando a leitura, cheia de suspense e aprendizado.

Referências

ANDERSON, Philip W. More is different. *Science*, v. 177(4047), p.393-396, 1973.

ENNS, Peter. The real problem evangelicals have with evolution (and what needs to be done about it), *Peter Enns blog*, 2012. Disponível em: <peteenns.com/deep-problem-evangelicalism-evolution>. Acesso em: 26 abr. 2020.

MARIN, Ronaldo; AMARAL, Gustavo Rick. *Os avanços da ciência ponderiam acabar com a filosofia?* São Paulo: Estação das Letras e Cores, 2020.

SCHULZ, Peter. A celebridade e o todo. *Comciência*, dossiê 197, 2018. Disponível em: <comciencia.br/o-todo-e-celebridade>. Acesso em: 26 abr. 2020.

WOLCHOVER, Nathalie. A Fight for the Soul of Science. *QuantaMagazine*. Nova York: Simons Foundation. Dez, 2015. Disponível em: <quantamagazine.org/physicists-and-philosophers-debate-the-boundaries-of-science-20151216>. Acesso em: 26 abr. 2020.

SCHREIBMAN, Susan; SIEMENS, Ray; UNSWORTH, John (eds.). *A companion to digital humanities*. Oxford: Blackwell, 2004.

[dx.doi.org/
10.23925/1984-3585.2020i21p218-226](https://dx.doi.org/10.23925/1984-3585.2020i21p218-226)

Resenha do livro *A companion to digital humanities*, de Susan Schreibman, Ray Siemens e John Unsworth (eds.).

Victor Sancassani¹

A publicação em 2004 do livro *A companion to digital humanities* foi um marco na emergência das Humanidades Digitais (HD) tanto como um campo de estudos *per se*, e também como o termo que passaria a ser utilizado para designá-lo. Em uma conversa por e-mail com Matthew G. Kirschenbaum, John Unsworth relata que

A verdadeira origem desse termo [Humanidades Digitais] estava em discussão com Andrew McNeillie, o editor comissionado original de *Companion to digital humanities* da Blackwell. Começamos a conversar com ele sobre o projeto do livro em 2001, em abril e, ao fim de novembro, nós escalamos os colaboradores e discutimos o título do contrato. Ray [Siemens] queria “A companion to humanities computing”, visto que esse era o termo comumente utilizado naquela época; os membros editoriais e do marketing da Blackwell queriam “Companion to digitized humanities”. Eu sugeri “Companion to digital humanities” para afastar a ênfase na mera digitalização. (apud KIRSCHENBAUM, 2010, p. 56-57)

O surgimento desse campo específico de estudos se deu tendo em vista o crescimento da utilização de métodos computacionais nas ciências humanas e sociais que, excedendo os limites do campo tecnológico da computação e da informática, aos quais se restringiam em grande parte até a década de 90, se expandiu para os campos social, cultural e artístico. Suas aplicações abrangem desde a manipulação, processamento e análise de dados massivos (*big data*), captura de imagens, registro, reconstrução e preservação de objetos históricos, até a criação de novos objetos em forma digital, que cresceram e se desenvolveram com a criação de organizações e associações, centros e institutos, além de conferências, simpósios, periódicos, manifestos e livros voltados para o tema. Conforme a definição de Busa (2004, p. xvi), Humanidades computacionais, isto é, Humanidades Digitais, são “precisamente a automação de qualquer análise da ex-

¹ Mestre em Comunicação e Semiótica pela PUC-SP (2018).

cv Lattes: lattes.cnpq.br/1080828187286313. E-mail: vsancassani@gmail.com.

pressão humana possível (portanto, ela é primorosamente uma atividade ‘humanística’), no sentido mais amplo da palavra, da música ao cinema, do design e pintura à fonética”.

Nascida, portanto, na inter-, multi- e transdisciplinaridade, da união entre o tradicionalismo dos estudos e a usabilidade das novas tecnologias, segundo McCarty (2015) as Humanidades Digitais não compreendem somente uma mera atividade tecnológica impulsionada pela sua própria progressão evolucionária no tempo, nem mesmo um surgimento repentino de uma atividade singular, mas a capacidade de inventar e ser reinventado pela tecnologia. Tal atividade ainda pode ser encarada como uma nova heurística para os acadêmicos das Humanidades, permitindo observar relações que demandariam muito esforço ou até mesmo seriam improváveis de serem apontadas apenas pela dedicação da capacidade pessoal. Tendo por base o modelo temporal de Friedrich Strass de 1804, McCarty divide o desenvolvimento do campo em três fases baseadas na compreensão do início da Web como um dilúvio que se alastra: 1) Fase Antediluviana (1949-1991), iniciada por Roberto Busa, cujo trabalho pode ser considerado um dos, senão o pioneiro da utilização dos métodos computacionais para as pesquisas nas Humanidades, ainda na década de 50, até a criação da Web, com um domínio no campo linguístico e literário; 2) Fase Diluviana (1990-2004/5), centrado na Web, com a inserção de objetos em forma virtual e a exploração das capacidades da plataforma; e 3) Pós-Diluviana (2004/5-atualidade), quando as questões a respeito da visualização, localização, popularidade, redes sociais e a problematização disciplinar entram em campo.

Apesar de uma origem e um evidente foco na utilização dos métodos computacionais na manipulação, processamento e análise de textos, tal como é observado em grande parte dos textos da primeira edição em 2004, conforme o avanço tecnológico permitiu, novos ramos e campos do conhecimento usufruíram e ainda usufruem de tais métodos, alcançando inúmeras outras formas midiáticas, que ficaram mais evidentes nesta segunda edição. É nesse contexto que foi publicada, em 2016, a segunda edição revisada e ampliada, chamada *A new companion to digital humanities*, também editada por Susan Schreibman, Ray Siemens e John Unsworth, que conta com 56 autores – sendo 11 destes já conhecidos da primeira edição – e um total de 37 capítulos, distribuídos em cinco seções.

Enquanto na primeira edição o objetivo estava voltado mais para a apresentação do histórico do campo das Humanidades Digitais, com um foco discursivo nas pesquisas de editoração, arquivamento e análise textual ainda remanescentes das primeiras contribuições sobre o tema, mas

que também as ultrapassam em determinados aspectos, nessa segunda edição a abordagem firma-se, de fato, como uma heurística que excede o campo da prática tecnológica para o campo epistêmico. Tal fato é observado, inclusive, na diferença da divisão de seções adotada entre as duas edições. Na primeira, os textos se dividiam em quatro temas fundamentais: 1) História; 2) Princípios; 3) Aplicações e 4) Produção, Disseminação e Arquivamento. Na segunda edição, as contribuições foram organizadas sob cinco principais temas: 1) Infraestruturas; 2) Criação; 3) Análise; 4) Disseminação e 5) Passado, Presente, Futuro nas Humanidades Digitais. Vale notar que essa perspectiva é também enfatizada na distribuição desses mesmos títulos, que se desenvolvem a partir das questões basilares, sejam históricas e/ou metodológicas, passando ao nível da aplicabilidade, chegando à reflexão sobre a contribuição e as perspectivas futuras das Humanidades Digitais.

A seguir, apresentamos uma súpula da temática de cada seção e dos assuntos discutidos em seus respectivos capítulos, tendo em vista a variedade de temas percorridos ao longo do livro por cada autor ou grupo de autores, e das relações entre eles, que podem ser de interesse específico do público.

Na primeira seção, intitulada *Infraestruturas*, os quatro primeiros capítulos do livro focam justamente na sutil função basilar que os componentes de qualquer maquinário cumprem, mesmo quando imperceptíveis, aplicada às Humanidades Digitais. Ou seja, a estruturação de novos meios de captação e abordagem de fenômenos, surgidos da interseção entre pensamentos ditos puramente humanos e operações ditas puramente maquinais, e que exigem não somente uma reeducação, mas métodos adequados para a sua compreensão.

No capítulo 1, Jentery Sayers, Devon Elliott, Kari Kraus, Bethany Nowviskie e William J. Turkel exploram a intersecção dos mundos físico e virtual, por meio dos processos de computação física e da fabricação de *desktops*, onde *bits* e átomos compõem um espaço propício para o borbulhar da especulação criativa e a realização das conjecturas críticas. Conforme Sarah Kenderdine (capítulo 2), esse novo espaço para a criação também é o local que permite a emergência da visualização de ambientes imersivos e interativos como um “novo capítulo na História da Arte”, em que o espectador se torna um agente no meio digital, que afeta e é afetado pelo mundo, assim alterando todas as relações pré-concebidas de *corporalidade* e *performance*, *experiência compartilhada* e *materialidade*. É nesse mesmo sentido que Finn Arne Jørgensen (capítulo 3) examina a Internet

das Coisas num limite turvo entre o físico e o digital, como um sistema que não foca apenas na interconectividade e implementação de aparatos tecnológicos suprimidos das relações sociais – como por exemplo, na célebre ideia da “geladeira inteligente” –, mas também de uma complexidade sociotécnica que altera os modos de percepção das informações que fluem entre corpos, físicos e digitais e, conseqüentemente, afetam componentes sociais, culturais e políticos. Jennifer Edmond (capítulo 4) faz uma análise crítica das HD a partir dessas mesmas tensões, que não somente fazem parte do mundo externo, mas também do próprio ecossistema acadêmico. As distintas disciplinas que são colocadas em colaboração também geram conflitos entre as diferentes perspectivas intelectuais, demandando a gerência de projetos, suas equipes e seus usuários, que parte desde o domínio ético e moral, da autoria de materiais e dados, ao processo de uma educação e aprendizagem eficientes.

Na segunda seção, intitulada *Criação*, essas ideias sobre a necessidade de uma reeducação ficam mais claras com as discussões das questões interdisciplinares inerentes levantadas a respeito das Humanidades Digitais, que colocam as pessoas como contribuidoras nesses novos processos. Os usuários passam da mera atividade receptiva de informação e de doação de dados para a de agentes criadores nas plataformas digitais, cujos *feedbacks* são valiosos para o aperfeiçoamento do próprio meio digital.

Willard McCarty (capítulo 5) trata especialmente das estratégias de adequação à vastidão de possibilidades emergentes do campo, advindas não somente do choque entre diferentes disciplinas, mas da própria natureza indagadora humana, que devem ser orientadas por uma meta-disciplina. Nessa mesma linha, tendo em vista a demanda por uma educação computacional que exceda os limites instrumentalistas das práticas do desenvolvimento de softwares, Nick Montfort (capítulo 7) concebe o conceito de *educação exploratória*, uma educação investigativa focada na construção de um pensamento crítico para a resolução de problemas que não levem somente a novas maneiras de pensar, tal como pode ocorrer a um indivíduo ou um artista, mas também que levem a novas maneiras de entender e contribuir na sociedade como um todo. Steven E. Jones (capítulo 6) aponta para a necessidade do estudo dos jogos digitais sob a perspectiva das Humanidades Digitais, como um sistema criativo que combina a materialidade da computação com os modos de expressão da dinâmica sociocultural, simulando a relação entre humanos e computadores, ou seja, para além do estigma de um mero entretenimento de massa. Do mesmo modo, Christopher Johanson (capítulo 8) amplia essa visão para a

criação de mundos virtuais como um todo, em que aquele que vivencia a experiência, além de interagir, possa também ter um ato demiúrgico e ser capaz de criar o mundo à sua própria imagem. Tal capacidade de contribuição da sociedade também é apontada por Kenneth M. Price (capítulo 10) no conceito de *editoração social*, cujos usuários das edições eletrônicas passam a ser colaboradores lado-a-lado dos profissionais, consequentemente gerando questionamentos tanto sobre a atuação, como sobre os modos apropriados de realizar tal troca de ideias e informações. Tendo em vista que diferentes grupos de usuários interagem com diferentes tipos de sistemas, esse mesmo contexto levanta a questão colocada por Séamus Lawless, Owen Conlan e Cormac Hampson (capítulo 12) sobre a importância da criação de novas técnicas de personalização e tecnologias de contextualização adaptativas para a apresentação e o acesso ao conteúdo digital, a fim de fornecer uma melhor experiência para os usuários. Lorna Hughes, Panos Constantopoulos e Costis Dallas (capítulo 11) desenvolvem uma ontologia dos métodos digitais aplicados às Humanidades Digitais no ecossistema acadêmico, que aborda tanto a proliferação, disseminação e colaboração interdisciplinar do conteúdo digital, como também apelos para uma melhor articulação e definição do conceito (2016, p. 151). Scott Rettberg (capítulo 9) se ampara na perspectiva do surgimento de novos campos do conhecimento, em defesa de uma literatura eletrônica que vai além da aplicação de ferramentas digitais às formas tradicionais de texto, alcançando o status de um campo *per se*. Por fim, no último texto desta seção sobre criação nas HD, Matthew G. Kirschenbaum (capítulo 13) explora na retrocomputação uma oportunidade de oferecimento de ferramentas e métodos para os questionamentos das HD, de modo a formular o exame crítico do legado tecnológico e o seu consequente desenvolvimento.

Na terceira e mais extensa seção do livro, intitulada *Análise*, inúmeros estudos que permeiam diferentes formas midiáticas são abordados nas suas qualidades práticas e como objetos de estudos do campo das Humanidades Digitais. No entanto, tais capítulos não têm por questão a habitual marca da técnica em si, mas principalmente de sua construção e seus efeitos aplicados adequadamente no meio, a fim de proporcionar melhores experiências e resultados para aqueles que as utilizam, seja na esfera cotidiana ou acadêmica.

Todd Presner e David Shepard (capítulo 14) exploram o campo da visualização geo-temporal (GPS e GIS) como um local crítico e narrativo da integração de diversos aspectos culturais que são mapeados pelas práticas das Humanidades Digitais. John Ashley Burgoyne, Ichiro Fujinaga e J. Stephen Downie (capítulo 15) apresentam as abordagens mais recentes a

respeito dos estudos de MIR e como dados visuais, sonoros e metadados são utilizados para a sua compreensão. Julia Flanders e Fotis Jannidis (capítulo 16) apresentam um estudo do impacto e influência dos métodos utilizados pelas Humanidades Digitais na modelagem de dados, que exige uma ação autorreflexiva de ambos os lados. Johanna Drucker (capítulo 17) avalia o ato de repensar os modos de apresentação das visualizações e interfaces digitais, que carecem dos princípios de estruturação e necessitam de uma perspectiva humanística-crítica para a sua adaptação gráfica, tornando-as mais eficientes. Dominic Oldman, Martin Doerr e Stefan Gradmann (capítulo 18) exploram o campo da Web Semântica e seus *linked data*, como um par complementar das conexões realizadas entre os dados da Web, de modo a construir uma rede que ultrapasse o âmbito meramente informacional, como uma fonte de significados profundos para futuros estudos. Visualização de dados também é o objeto de estudo de Stéphan Sinclair e Geoffrey Rockwell (capítulo 19), que as identificam como práticas interpretativas com um alto nível de produção de significado, desde que trabalhada cuidadosamente para que conexões notáveis sejam percebidas. Apesar de um enfoque literário, Matthew L. Jockers e Ted Underwood (capítulo 20) enfatizam o poder sintético da mineração de dados como uma ferramenta sofisticada e promissora para estudos em diversas disciplinas das Humanidades. Seguindo na linha textual, Elena Pierazzo (capítulo 21) examina a codificação textual como uma mudança heurística no modo de pensar os estudos digitais, cujo trabalho editorial contribui para uma readaptação contextual, de modo a satisfazer adequadamente as necessidades de seu público. Por sua vez, Sydney J. Shep (capítulo 22) foca no conceito de materialidade digital como um elemento quase invisível, mas cujos vestígios são essenciais como evidências para a compreensão “dos conceitos, métodos e práticas nas e das Humanidades Digitais” (2016, p. 328). Joris J. van Zundert (capítulo 23) atenta para a importância de considerar computadores como instrumentos hermenêuticos, ou seja, não somente como produtores de códigos, quantificações e algoritmos, mas a sua extensão significativa para os estudos nas Humanidades. Tanya E. Clement (capítulo 24) apresenta uma perspectiva crítica sobre a carência de acesso e uso de materiais sonoros nas pesquisas em HD, em contraponto aos materiais visuais e textuais, devido principalmente, tanto à importância do significado do conteúdo como dos métodos para sua obtenção. Retornando às questões textuais, Jerome McGann (capítulo 25) expõe a utilização de ferramentas de processamento de textos para o seu mapeamento e consequentes revisões,

contribuições e comparações críticas, de modo a torná-lo um evento auto-poético. C. M. Sperberg-McQueen (capítulo 26) finaliza a seção analítica da obra ressaltando a necessidade da criação de sistemas classificatórios para a organização e gerenciamento das informações coletadas, de acordo com as estruturas que os compõem.

Na quarta seção, intitulada *Disseminação*, a questão das Humanidades Digitais chega efetivamente às pessoas, tanto aos usuários do meio digital como da sociedade em geral, mesmo que seus membros sejam afetados indiretamente. Os artigos desta seção ressaltam a importância do agenciamento dos usuários na construção do ambiente digital e o papel intermediador da tecnologia nessa interação.

Stan Ruecker (capítulo 27) apresenta um estudo sobre as interfaces à luz da teoria ator-rede, cujos atores passam a ser, não apenas os usuários, mas também a própria interface, ao fortalecer a intermediação realizada entre as pessoas e a tecnologia. No ambiente digital tudo se altera e é substituído muito rapidamente. E é nesse contexto que William Killbride (capítulo 28) atenta para a questão da preservação digital, que sobrepassa a esfera prática da perda de dados ao inserir o pensamento preventivo nas próprias estruturas digitais, uma vez que dados tratam de pessoas e oportunidades. Dando continuidade, Melissa Terras (capítulo 29) também se preocupa com a utilização e reutilização dos dados, porém com foco na inserção e no engajamento do público nas pesquisas humanísticas, integrando setores não acadêmicos à área, de modo a beneficiar ambos os lados na compreensão da sociedade e de nossa herança cultural. Tendo em vista o aumento das publicações em periódicos digitais, Kathleen Fitzpatrick (capítulo 30) explora novos métodos de avaliação por pares que atendam às novas formas de criação de textos nativos na Web, cuja interpretação e comunicação de dados sejam não somente acessíveis, mas também legíveis para membros de outras comunidades científicas. Todas essas questões se interseccionam no texto de Stephen Ramsay (capítulo 31). O autor retira a atenção da palavra “mídia” para ressaltar a importância dos atores humanos no meio digital, como aqueles que estão criando, aprendendo, interagindo, reconfigurando e sendo reconfigurados.

A quinta e última seção, *Passado, Presente, Futuro nas Humanidades Digitais*, trata das perspectivas futuras das HD, daquilo que ainda precisa ser realizado e questionado no que concerne ao seu impacto na sociedade, seja nos modos como elas atingem e a quem atingem, mas principalmente como torná-las mais presentes e incorporadas no meio social de maneira a favorecer todos os membros igualmente.

Andrew Prescott (capítulo 32) explora a importância dos centros de Humanidades Digitais como veículos institucionais que tanto transformam os modelos tradicionais de pesquisas acadêmicas, como criam uma rede interativa e colaborativa com os profissionais de dentro e de fora da universidade, e entre pessoas de diferentes partes do mundo. Patrik Svensson (capítulo 33) busca revitalizar a questão do papel das HD como ponto intermediário entre os diferentes campos do conhecimento, ao sair em defesa do fator liminar e não-territorial do campo, visto como uma vantagem disciplinar para o seu desenvolvimento, ao invés de um entrave. Daniel Paul O'Donnell, Katherine L. Walter, Alex Gil e Neil Fraistat (capítulo 34) apresentam uma reflexão sobre a expansão dos estudos das HD para além das estruturas e pessoas. Nele, os autores evidenciam a necessidade de reconhecimento, validação e acomodação de fatores como as barreiras geográficas, culturais, econômicas e linguísticas, que precisam ser cruzadas a fim de garantir melhores modos de estabelecer redes conectivas entre diferentes partes do mundo. Laura C. Mandell (capítulo 35) levanta a questão do ofuscamento das escritoras na imensidão de dados do meio digital como um reflexo da própria exclusão no cenário tradicional literário impresso, em prol da importância de uma alavancagem da história literária digital feminista. William G. Thomas III (capítulo 36) ressalta a necessidade de construir novas maneiras de pensar os estudos digitais dentro das Humanidades Digitais, de modo que elas criem um ambiente de discussão sobre a influência do meio digital em diversas disciplinas e nas práticas institucionais. Claire Warwick (capítulo 37) finaliza o livro apresentando as divergências existentes nas Humanidades Digitais entre os seus aspectos práticos e teóricos, mostrando-se também como um campo fértil para o debate da sua constituição e aplicabilidade, em vias de reivindicação de um campo de estudo único.

A diversidade de temas discutidos em *A new companion to digital humanities* atualiza e ilustra a vastidão e a crescente relevância das HD como interface expansiva que possibilita planejamento e execução de projetos inter-, multi- e transdisciplinares.

Referências

BUSA, Roberto. Foreword: perspectives on the digital humanities. In: SCHREIBMAN, Susan; SIEMENS, Ray; UNSWORTH, John (eds.). *A companion to digital humanities*. Oxford: Blackwell, p. xvi-xxii. 2004.

KIRSCHENBAUM, Matthew G. What is digital humanities and what's it doing in English departments? *Association of Departments of English: ADE Bulletin*, n. 150, 2010, p. 55-61.

MCCARTY, Willard. The big picture: where digital humanities has been & where you might take it. Krasnoyarsk: Siberian Federal University, 25 de setembro de 2015 (=Workshop Digital Humanities Conference). Disponível em: <mccarty.org.uk>. Acesso em: 23 mar. 2020.

SCHREIBMAN, Susan; SIEMENS, Ray; UNSWORTH, John (eds.). *A companion to digital humanities*. Oxford: Blackwell, 2004.

SCHREIBMAN, Susan; SIEMENS, Ray; UNSWORTH, John (eds.). *A new companion to digital humanities*. Chichester/Malden: Wiley/Blackwell, 2016.

BERRY, David M.; FAGERJORD, Anders. *Digital humanities: knowledge and critique in a digital age*. Cambridge, UK/Malden, USA: Polity, 2017.

Resenha do livro **Digital humanities: knowledge and critique in a digital age**, de David M. Berry e Anders Fagerjord

Monica Pinheiro Allan¹

Berry e Fagerjord se propõem a debater as genealogias das Humanidades Digitais (HD) a partir da transição do *computing in the humanities* para as chamadas *digital humanities*. O livro desdobra-se em argumentações dinâmicas a partir da multiplicidade de vozes autorais que irão constituir o mapeamento, a fim de fomentar a criticidade epistêmica e metodológica acerca das HD, mas também em relação às questões práticas do campo e ao desenvolvimento do ferramental da computação.

Berry e Fagerjord exemplificam como a cultura digital surge a partir das mediações de hábitos e das exigências formais advindas da complementaridade entre o analógico e o digital, o que, às vezes, torna “aquilo que é digital” secundário frente a muitas formas de mediação, produção, acesso, distribuição e consumo através de dispositivos digitais e tecnológicos. As HD situam-se como um campo de pesquisa novo e o livro se posiciona em meio aos debates da tecnologia e cultura. Os autores ressaltam o momento marcadamente histórico em que escrevem, visto que as transformações criativas no meio acadêmico e no modo social tornam a produção digitalizada de conhecimento em rede contínua nos intermeios da conexão e desconexão. A ubiquidade da tecnologia frente à sociedade pós-digital exige um preparo para futuras intervenções críticas, seja sobre o gerenciamento do conhecimento técnico e seus agenciamentos ou sobre a discussão da ética em relação às diversas formas de tecnologias e seus usos, nem sempre visíveis a todos. O livro, em sua orquestração heurística sobre cultura e conhecimento, apresenta timbres e inflexões de reverses polêmicos, até antitecnológicos nas bifurcações dos interesses humanos, como modo de ajudar a formar maior criticidade em relação ao tema. Para isso, Berry e Fagerjord destacam autores consagrados em discussões textuais e filosóficas, tais como Schreibman, Pannapacker e Moretti para iluminarem a temática e seus desdobramentos.

¹ Mestre em Comunicação e Semiótica pela PUC-SP (2018).

cv Lattes: lattes.cnpq.br/0730011721582342. E-mail: milvezesmonica@gmail.com.

Na Introdução, Berry e Fagerjord apresentam as *Humanidades Digitais*, ou simplesmente HD, como termo que ainda traz em si muitos embates intelectuais a seu respeito, e que se encontra marcado na sua origem engendrada no movimento que ficou conhecido como *humanidades computacionais*. Para os pensadores dessa linha de pesquisa, trata-se de uma disciplina na qual há muito a se construir, uma vez que emergiu como movimento sob essa denominação no âmbito acadêmico somente a partir de 2001. De maneira geral, sob o termo HD, é possível encontrar pesquisas especializadas, demarcadas por necessidades específicas do meio em que foram desenvolvidas, orientadas pela lucidez da consciência que, ao mesmo tempo, assiste e compartilha as potencialidades do digital nas universidades e que é cônica da importância de incorporar e corporificar as experiências e modelos digitais como base fundamental para as humanidades. Não sem crítica, e nem sempre com a visibilidade dos efeitos que essa transformação codificada causa através de um mundo de acessos, de superfícies, de redes, de dados (*big data* e *meta data*), de arquivos e imagens, logaritmos, dispositivos e interfaces, e de todo um ferramental tecnológico que somente o pensamento rizomático pode alcançar em seus múltiplos campos de uso e inventividades.

Berry e Fagerjord apresentam um livro guia, cujos argumentos teóricos são apresentados de maneira a articular uma virada crítica do pensamento sobre as humanidades, contribuindo para o desenvolvimento de uma importante perspectiva cognitiva. Os autores partem do modelo da comunicação midiática, das métricas entre educação e consumo para inserirem as humanidades no contexto digital e discutir se elas podem nos levar ao futuro. Grusin, sobre o lado escuro e o lado luminoso das HD, Hayles, McGann, Hitchcock e Kirschenbaum são alguns dos teóricos ilustrados nesse mosaico que se aprofunda nas preocupações tradicionais da produção do conhecimento e em suas representações por meio de métodos, infraestrutura de pesquisa e armazenamento digitais. Drucker é lembrado (p. 12), assim como opiniões controversas como forma de direcionar os sentidos para estudos humanísticos sob os efeitos da tecnologia, na busca de uma teoria da informação humanística para compor o uso da tecnologia, que estruture a arquitetura da modelagem dessa informação presente nas leituras da mídia social e em jogos, narrativas, tipos de dados, interfaces e protocolos.

Na busca por um diagrama, os autores ilustram de modo gráfico a abstração criativa da informação em camadas, a qual chamam de “pilha das HD” (p. 19). Linguagem comum na computação, esse recurso é usado no livro para mostrar como as tecnologias formam camadas, servindo ao propósito ilustrativo da busca de metáforas para ajudar na compreensão da complexidade criativa do digital.

No capítulo 2, “Genealogies of digital humanities”, Berry e Fagerjord desenvolvem as origens do *computing in the humanities* (p. 25). O campo das HD é explorado conceitualmente. Roberto Busa (1913-2011) é um dos pioneiros, associado ao início do uso da computação nas humanidades. Padre jesuíta e professor de filosofia, ele se destacou nas humanidades computacionais ao indexar a obra de Tomás de Aquino, usando um computador IBM em 1949. As habilidades necessárias ao trabalho no século XXI compõem debates teóricos sobre o trabalho digital em seu paradigma de função, de cooperação e de conhecimento aprofundado como benefícios trazidos pelas HD, bem como sobre sua complexidade, que abarca vários campos como educação, política, sociologia, além da própria computação e de diversas outras linguagens (linguística, metalinguística e metalinguagem).

Questões sobre a automação, código aberto (*open source*) e cocriação são trazidas à tona, bem como discussões importantes sobre o que são as HD em si e quais as práticas ferramentais criativas adotadas. A cultura “industrializada” toma para si as descobertas científicas? Quais são os modelos adotados pelas HD? Questões importantes como essas são debatidas. O livro apresenta outras questões sobre documentos abertos e interfaces colaborativas que pressupõem as leituras de Cathy Davidson (2012), *Humanities 2.0*, e de Presner e Schnapp (2009), *Manifesto 2.0*, publicado pela Universidade da Califórnia. A chave para a argumentação de Berry e Fagerjord é a de que as disciplinas se fortalecem tendo pesquisadores especializados e que, portanto, podem se mover entre projetos de HD e teorias (p. 34). Berry (p. 35), que dividiu analiticamente as HD em 1ª Onda: *Computing in humanities* (1940-2001); 2ª Onda: *Digital humanities* (HD) – interfaces e nascido digital (*born digital*), 3ª Onda – *Digital humanities*, materialidade e cultura crítica, nesse capítulo também apresenta outros modelos, inclusive o Tipo 1 e Tipo 2 de HD de Ramsay (p. 36).

No capítulo 3, “On the way to computational thinking”, os autores apresentam o pensamento computacional (p. 41-42) e um panorama dos elementos comuns a esse campo rumo às HD. Os autores acreditam ser fundamental à fomentação crítica compreender as chaves computacionais: automação, abstração, decomposição, algoritmos, aprendizado das linguagens e estética da computação. Através da exemplificação, também discutem processo e modelagem para abordarem a questão humanista através da problematização e inserção de processos na busca de soluções. Destacam sempre a importância das pesquisas, ilustram a relevância das diferentes formas de Inteligência Artificial (p. 50) e seu uso em plataformas como Google Search, Amazon e Flickr, e a importância econômica

de corporações midiáticas como Apple, Facebook etc. na sistematização de algoritmos para desenvolver a programação, o insight criativo e a representação computacional para as HD.

O capítulo 4, “Knowledge representation and archives”, desenvolve as abordagens a respeito de metadados em suas diferentes representações. Os autores estão preocupados com a formatação material da cultura em imagens, livros, artigos, sons, filmes, vídeos etc. Assim como Flusser, que versa sobre artefato e artifício, Berry e Fagerjord estão sob o mote de Minsky (p. 61) quando escreveu que “o modelo não é o objeto em si, mas a representação daqueles aspectos do objeto que preocupam os pesquisadores”. Os autores abordam diversas linguagens (p. 65), entre elas SGML (*Standard Generalized Markup Language*), XML (*Extensible Markup Language*), GIS (*Geographic Information System*), investigam os métodos e técnicas utilizadas, e questionam as linguagens que logo estarão muito além do que é compatível com a epistemologia tradicional e as disciplinas humanistas (p. 75-79).

O quinto e sexto capítulos são a extensão dessa lógica investigativa sobre a computação, construindo os caminhos tecnológicos necessários às HD. No capítulo 5, “Research infrastructures”, Berry e Fagerjord “querem gastar mais tempo pensando a respeito dos caminhos pelos quais as infraestruturas apoiam a pesquisa e a tornam possível como ensino nas universidades” (p. 80). No capítulo 6, como o título já diz, “Digital methods and tools”, trabalha a construção dos sistemas digitais (p. 103). Para os autores, as HD, não apenas consideradas como construção, são frequentemente situadas em diferentes tipos de estruturas e instituições, o que faz delas campos cognatos nas humanidades. Arquivo, processo, software, interfaces públicas e codificação são estudados sob o axioma cunhado por Richards Rogers: métodos digitais para aquilo que já nasce digital (p. 108-109).

O capítulo 7, “Digital scholarship and interface criticism”, e o capítulo 8, “Towards a critical digital humanities”, levam o leitor à criticidade (p. 114) proposta pelos autores no subtítulo do livro e apresentada como objetivo logo na Introdução. Aqui, Berry e Fagerjord edificam sua base epistemológica e traçam o arco da ética nas funções digitais, na construção crítica fundamentada e argumentada sob vários prismas. Não será o computador a verdade absoluta enquanto resposta, porque são as perguntas sobre tudo aquilo que não está visível nas telas que importam. As camadas interfaceadas, a transparência, a aplicação do *Big Data* (p. 137), a cultura mediando a própria cultura. A essa proposta especulativa aberta e intensa, principalmente na conclusão do livro, os autores evocam e

nomeiam como *Humanidades Digitais críticas* (138-140), porque elas mesmas se entretecem nas questões de como o conhecimento se transformou quando foi mediado pelo código e pelo software. O livro em si é um documento sobre como as HD têm crescido e desenvolvido suas potencialidades para futuras possibilidades, em suas palavras autorais.

Referências

- BERRY, David M. The computational turn: thinking about the digital humanities. *Culture Machine*, 12. Open Humanities Press, 2011. Disponível em: <n9.cl/pdmt>. Acesso em: 31 ago. 2020.
- BERRY, David M.; FAGERJORD, Anders. *Digital humanities: knowledge and critique in a digital age*. Cambridge, UK/Malden, USA: Polity, 2017.
- DAVIDSON, Cathy N. Humanities 2.0: promise, perils, predictions. *Debates in the Digital humanities*. Matthew K. Gold (ed.). Minneapolis: University of Minnesota Press, 2012. Disponível em: <n9.cl/1bv78>. Acesso em: 31 ago. 2020.
- HAYLES, N. K. How We Think: Transforming power and digital technologies. *Understanding the digital humanities*. D. M. Berry (ed.). London: Palgrave, 2011.
- KIRSCHENBAUM, Matthew. What is 'digital humanities', and why are they saying such terrible things about it? *Differences*, n. 25, vol. 1. Elizabeth Weed & Ellen Rooney (eds.). Duke University Press, 2014, p. 46-63 Author's PDF. Disponível em: <n9.cl/tvfi>. Acesso em: 31 ago. 2020.
- LIU, Alan. Where is cultural criticism in the digital humanities? *Debates in the digital humanities*. Matthew K. Gold (ed.). University of Minnesota Press, 2012, p. 490-509. Disponível em: <n9.cl/1bv78>. Acesso em: 31 ago. 2020.
- SCHNAPP, Jeffrey; PRESNER, Todd. Digital humanities manifesto 2.0. University of California, Los Angeles, *UCLA Website*, postado em 29/05/2009. Disponível em: <manifesto.humanities.ucla.edu/2009/05/29/the-digital-humanities-manifesto-20>. Acesso em: 07 set. 2020.



Diretrizes para autores – TECCOGS

A *TECCOGS – revista digital de tecnologias cognitivas* é um periódico do Programa de Pós-Graduação em Tecnologias da Inteligência e Design Digital (TIDD) da Pontifícia Universidade Católica de São Paulo (PUC-SP). As edições são semestrais e exclusivamente digitais, disponíveis em pucsp.br/pos/tidd/teccogs.

A **TECCOGS recebe artigos e resenhas de doutores ou de especialistas, mestrandos, mestres e doutorandos em coautoria com doutores.**

Título, subtítulo, resumo (com no mínimo 1000 e no máximo 2500 caracteres com espaços) e **palavras-chave** (de três a seis termos) do artigo deve aparecer em português ou espanhol (caso o artigo esteja escrito nessa língua) e, logo em seguida, traduzidos para o inglês.

O(s) **nome(s) do(s) autor(es)** deve(m) estar logo abaixo do subtítulo do artigo, acompanhado de uma nota de rodapé (escrita em fonte *Times New Roman* tamanho 11 pt, espaçamento simples) contendo currículo e biografia (formação, vínculo acadêmico, área de atuação e e-mail) com, no máximo, cinco linhas.

Cada artigo deve possuir no mínimo 20.000 e no máximo 50.000 caracteres com espaços.

Resenhas devem possuir no mínimo 8.000 e no máximo 13.000 caracteres com espaços.

O **corpo do texto** deve ser configurado em fonte *Times New Roman* tamanho 12 pt, espaçamento 1,5 linhas, parágrafo alinhado à esquerda, sem hifenização. **Citações diretas com quatro linhas ou menos** devem aparecer entre aspas (“”) incorporadas ao corpo do texto, indicando a fonte entre parênteses no modelo “(SOBRENOME [em maiúsculas], ano de publicação, p. [número da página])”, conforme a [Norma Brasileira \(NBR\) 10520 \(ago. 2002\) da ABNT](#).

As **citações diretas com mais de quatro linhas** devem ter recuo à esquerda de 4 cm, sem aspas, com fonte *Times New Roman* tamanho 11 pt, espaçamento simples, parágrafo justificado e sem hifenização.

Imagens (fotografias, ilustrações, diagramas, tabelas, gráficos) precisam ter resolução de, no mínimo, 100 dpi/ppi (*pixels* por polegada) e devem estar integrados ao corpo do texto, com imagem e legenda centralizadas e fonte especificada (para imagens da *internet*: “Disponível em: “<site>”. Acesso em: “dia mês abreviado ano”).

O texto deve respeitar o **Novo Acordo Ortográfico da língua portuguesa**, vigente desde 2009. De acordo com a [Base XIX da Nova Ortografia](#), termos como “Inteligência Artificial”, “Psicologia Cognitiva”, “Informática” e “Filosofia” (quando se trata da área de conhecimento) devem iniciar com maiúsculas. Segundo a [política de direitos autorais da revista](#), os autores se responsabilizam pelos direitos de uso de todas as imagens.

Para elaboração de resumos, citações e referências, a revista segue as NBR [6023 \(ago. 2002\)](#), [6028 \(nov. 2003\)](#) e [10520 \(ago. 2002\)](#) da ABNT. Não são permitidas notas de fim. Notas de rodapé devem ser usadas o mínimo possível, exclusivamente para adicionar observações pontuais, nunca para indicar referências bibliográficas. Em fontes da *internet*, a autoria do texto deve ser indicada entre parênteses, bem como o ano de publicação e endereço e data de acesso.

Todas as obras mencionadas nas referências devem estar citadas ao menos uma vez no texto e, do mesmo modo, toda e qualquer obra mencionada no texto deve constar nas referências.

A **TECCOGS** disponibiliza um arquivo formato .DOC que serve de *template* com instruções e exemplificações e estilos detalhados para escrever o artigo. [Baixe o modelo aqui](#).