

Objetos digitais: da maleabilidade do não-finito à uma ontologia

Objetos digitales: de la maleabilidad del no-finito a una ontologia

Digital objects: from malleability the non-finite to an ontology

Renan Marinho de Castro¹

Palavras-chave:

Digitalização

Acervos Digitalizados

Objetos Digitais

Humanidades Digitais

Ontologia do Digital

Resumo:

Apresenta os objetos digitais e seu impacto às unidades de informação utilizando a análise do caso do Centro de Pesquisa e Documentação de História Contemporânea do Brasil da Fundação Getulio Vargas (FGV/CPDOC). Destaca, brevemente, o impacto das tecnologias no cotidiano social contemporâneo, e sua influência nas transformações observadas no campo informacional. Manifesta a necessidade de uma reflexão, por parte das Unidades de Informação, a respeito do impacto que os documentos digitalizados e digitais têm sobre seu cerne, uma vez que com a criação de uma massificação do digital, se promove uma maleabilidade que suportes tradicionais de natureza física não comportam. Conceitua o *born-digital* e o destaca como sendo um elemento capaz de renovar toda a cadeia lógica informacional. Analisa os dados de consulta ao acervo do CPDOC, de acordo com as coleções mais visitadas no período de 2010 a 2016, destacando o impacto no número de acessos a partir da disponibilização de cópias digitais das coleções. Aponta a ontologia do digital como uma possibilidade de apropriação da informação e suas repercussões sob contexto da revolução tecnológica.

Resumen:

Se presenta los objetos digitales y su impacto a las unidades de información utilizando el análisis del caso del Centro de Investigación y Documentación de Historia Contemporánea de Brasil de la Fundación Getulio Vargas (FGV / CPDOC). Destaca, brevemente, el impacto de las tecnologías en el cotidiano social contemporáneo, y su influencia en las transformaciones observadas en el campo informacional. Se manifiesta la necesidad de una reflexión por parte de las Unidades de Información sobre el impacto que los documentos digitalizados y digitales tienen sobre su núcleo, ya que con la creación de una masificación de lo digital, se promueve una maleabilidad que soportes tradicionales de naturaleza física no comportan. Conceptua el born-digital y lo destaca como un elemento capaz de renovar toda la cadena lógica informacional. Analiza los datos de consulta al acervo del CPDOC, de acuerdo con las colecciones más visitadas en el período de 2010 a 2016, destacando el impacto en el número de accesos a partir de la disponibilización de copias digitales de las colecciones. Se apunta la ontología del digital como una posibilidad de apropiación de la información y sus repercusiones bajo el contexto de la revolución tecnológica.

Palabras clave:

Digitalización
Acervos Digitalizados
Objetos Digitales
Humanidades Digitales
Ontología del Digital

Keywords:

Digitalization
Digitalized Collections
Digital Objects
Digital Humanities
Humanities Computing
Digital Ontology

Abstract:

This text presents the digital objects and their impact to information units using the case study of the Getulio Vargas Foundation (FGV / CPDOC) Center for Research and Documentation of Contemporary History of Brazil. It briefly highlights the impact of technologies on the contemporary social, and its influence on the transformations observed in the informational field. It points out the need for a reflection on the impact of digitized and digital documents on their core by the Information Units, since with the creation of a mass of digital, it is promoted a malleability that traditional supports of physical nature do not support. It conceptualizes the born-digital and highlights it as being an element capable of renewing the entire informational logical chain. It analyzes the consultation data to the CPDOC collection, according to the most visited collections in the period from 2010 to 2016, highlighting the impact on the number of accesses from the availability of digital copies of the collections. It points to the ontology of digital as a possibility of appropriation of information and its repercussions in the context of the technological revolution.

Objetos digitais: da maleabilidade do não-finito à uma ontologia

Introdução

Mudanças contínuas e sem precedentes experimentadas nos meios social e natural da ação humana trouxeram em seu bojo experimentações que pavimentaram o caminho para conceitos que hoje balizam a sociedade contemporânea e permeiam os estudos sobre esse processo. Sendo esse movimento contemplado por diversos ângulos – econômico, cultural, político e social – fizeram surtir abordagens dedicadas a analisar esses fenômenos. De certo esse processo foi o responsável por construir temas maciçamente difundidos e conhecidos atualmente como, por exemplo, Globalização, Sociedade em Rede, Sociedade da Informação, Revolução Tecnológica, Era da Informação, entre tantos outros.

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

“centralidade de conhecimentos e informação, mas a aplicação desses conhecimentos e dispositivos de processamento da informação em um ciclo de realimentação cumulativo entre as inovações e seu uso” (CASTELLS, 2000, p.69).

Paralelo a essa abordagem temos a ideia de sociedade da informação, que articula esse movimento dentro de uma lógica na qual o processo não se caracteriza pelo mero ciclo de realimentação cumulativos entre as inovações e seus usos, como parece ser consagrado na sociedade em rede. Pela sociedade da informação, segundo Marques e Pinheiro (2013, p.117), “cria-se a necessidade do estabelecimento de novas políticas, que atendam a tais transformações”. Para as autoras destacam-se neste cenário discussões relativas ao acesso à informação e aos atores responsáveis pela função de democratizá-lo. Vê-se, a partir de então, um direcionamento dessas reflexões voltando-se às problemáticas do campo informacional ou este se apropriando daquelas.

No campo informacional essas mudanças se fizeram sentir ainda com maior voracidade. Das formas de compra e leitura de um livro à consulta ao catálogo da biblioteca, a tecnologia pautou autoritariamente a realidade de instituições de informação como bibliotecas, centros de documentação e afins. Esse processo, aliado à explosão informacional^{II}, que Varela (2005, p.2) pontua como o movimento em que destacam-se a abordagem da realidade complexa através de ferramentas cognitivas; a construção de representações mentais; o desenvolvimento de competências e habilidades; e a reflexão em torno da apreensão e da compreensão da informação pelo sujeito.

Nesse sentido, de certo, podemos cogitar que o digital vem promovendo uma revolução tal qual a industrial, por impactar a produção, o armazenamento, a recuperação, o acesso e a divulgação da informa-

ção e do conhecimento. Sua maleabilidade possibilita novos usos e apropriações elevando o patamar da relação homem *versus* informação. Todo esse caráter inédito até então, vem gerando reflexões e movimentos em torno do que se categoriza como revolução digital que tem se consolidado, especialmente no mundo corporativo, sob a alcunha de “Transformação Digital”. Para a Ross e Sennyey (2008) a revolução digital reformou todas as etapas de circulação da informação aotornar o formato predominante, alterou a velocidade de criação da informação, da entrega a divulgação, além das necessidades e expectativas dos consumidores de informação.

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

Da maleabilidade do digital a uma ontologia

O movimento de automatização de praticamente toda a interação humana teve na digitalização seu maior expoente, pois através desse processo de transformação pôde conferir diversas novas formas de transmissão, armazenamento e acesso a esses registros. Essa engrenagem, que anteriormente tinha sob sua gerência itens físicos por meio de computador como os sistemas de recuperação de

informações, hoje é capaz de produzir em ambiente totalmente não físico além de, também, igualmente gerenciá-lo. Assim, dessa forma, ao produto desse processo denominamos *Born-Digital*^{III}.

Como ‘*born-digital*’ entenda-se exclusivamente a informação produzida sob a égide do digital e sem ela inexistente. Como define Muhanna (2018) o *born-digital* significa, em alguns casos, referir-se a registros ou arquivos que não tem contrapartida física: coisas que começaram a existir enquanto entidades digitais em vez de prepostos análogos convertidos em forma digital (p.2, tradução nossa). Pelo que vimos até agora, é nesta entidade que se encontra não apenas a base para constituição de uma nova área - das digital humanities - a partir do avanço das reflexões propostas pela *Humanities Computing*^{IV}. O born-digital sinaliza a instauração de um elemento capaz de reordenar toda a cadeia da lógica informacional, uma vez que sua contribuição “não é apenas a introdução de um novo habitat” para o trabalho nas humanidades, mas sim o “oferecimento de um conjunto de novos mecanismos para obter, visualizar e manipular esses dados”. Ao fazê-lo, as tecnologias digitais contribuem com a dimensionalidade, eficiência, acessibilidade e interatividade para modelos estruturais que foram, por milênios, utilizados. (MUHANNA, 2018, p.1 tradução nossa).

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

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

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

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

De toda forma, vemos nesta imposição, o estabelecimento de um modelo que convergiu para a preponderância de um formato. Este modelo é abordado sob uma perspectiva filosófica por Hui (2002) que vê nesta instituição uma “ontologia do digital” que consiste em dois conceitos princi-

pais: primeiro, que os bits são a representação atômica do estado da informação; e, segundo, que o estado temporal da evolução é um processo de informação digital (p.381, tradução nossa). Encontramos nessa abordagem a fundamentação para considerarmos a importância deste tema independentemente de sua característica *born-digital*. Todavia sua segunda categorização deste conceito, em que atribui ao estado temporal da evolução enquanto um processo de informação, identifica, ao nosso ver, a questão da informação gerada exclusivamente em meio digital.

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

[...] devemos reconhecer que, desde 1946, a palavra “Dados” teve um significado adicional: “informação transmissível e armazenável por computador”. Este segundo sentido de “dados” sugere uma reconsideração da filosofia dos objetos, já que o dado não pode mais ser tomado como dados sensíveis ou um modo de estar junto do homem e da natureza; Em vez disso, é preciso reconhecer sua transformação material. O significado da nova técnica de processamento de dados que chamamos de digital não é apenas que com computadores podemos processar grandes quantidades de dados, mas também por, operando com dados, o sistema poder estabelecer conexões e formar uma rede de dados que se estende de plataforma em plataforma e banco de dados para base de dados. (HUI, 2002, p.388, tradução nossa)

Essa questão conduz a análise para a convergência das atuações dos indivíduos íntimos às computações e aqueles demais quais presenciaram a migração de seus objetos de análise para o digital.

Esse cruzamento, que tem em práticas de operação das informações digitais como XML, Web semântica, padrões de metadados etc., configura uma terceira fase. Para Hui (2002) a contraposição de obje-

tos naturais e digitais, o papel do 'dado' e o estabelecimento de redes fruto do processo do digital pontuam seu debate filosófico em torno dos objetos digitais dentro de 3 fases: *objects, data, and networks*.



Figura 1 – Fases na abordagem filosófica dos objetos digitais proposta por Hui (2002)

Fonte: DO AUTOR

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

Essas 3 fases, além de, ao nosso ver, credenciarem os objetos digitais a um patamar “não finito”,^{vi} podem encontrar reverberação em Yamaoka e Gauthier (2013) por, segundo estes, terem os objetos digitais três classes baseadas em Thibodeau (2002). Sendo essas classes, para eles, objetos de ordem física, lógica e conceitual. Em sua ordem física se aproxima da primeira fase de Hui (2002) que defende a resolução de objetos naturais ante os técnicos, cuja ideia predominante é sua qualidade física; em sua ordem lógica se aproxima de sua capacidade de reserva de ‘dados’, que conseqüentemente, se relaciona com

a ordem conceitual. Em sua terceira classe, produz, assim, sentido a partir dos dados e possibilitando então o estabelecimento de redes, que se configura enquanto última fase proposta por Hui (2002).

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

Por conseguinte as características do suporte digital foram responsáveis por pavimentar o caminho a uma nova estrutura de trabalho no âmbito das ciências, reprogramando totalmente os métodos e instru-

mentos de pesquisa, sobretudo nas disciplinas não íntimas, até então, ao fenômeno das TICs. Consequentemente esse processo, ao nosso ver, é parte indissociável ou mesmo condição *sinequa non* para o contexto dos ambientes digitais. Dessa forma subscrevemos o que Evens (2012) classifica enquanto uma ontologia do digital, que propõe uma reflexão a respeito do sentido abrangente deste atributo, como aquilo que torna possível suas múltiplas existências. Nos apropriando dessa concepção para pactuar através dela a transição para o que culminou na concepção de processos *dedicados à interpretação desse fenômeno no ambiente informacional como, por exemplo, as humanidades digitais*.

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

De certo consideramos que essa ontologia, expoente da predominância do instituto do digital, foi o ponto de partida, o gene por natureza, e cerne por consequência, do que promoveu transformações ímpares na práticas informacionais. Dessa forma responsabilizando-se por reunir em torno de uma incorporação abstração, pessoas, objetos, eventos e outros em co-

nexão mútua, influenciando seu comportamento e refletindo essa lógica. Assim revela-se um comportamento coletivo ante ao determinismo tecnológico que impulsiona a agregação de métodos, indivíduos e reflexões em torno das Humanidades Digitais.

As unidades de informação ante a ontologia de um novo suporte: a potencialização dos ambientes digitais

Para contextualizarmos o impacto que o digital imputa às unidades de informação recorreremos ao caso do Centro de Pesquisa e Documentação de História Contemporânea do Brasil da Fundação Getúlio Vargas (FGV/CPDOC)^{vii}. A instituição que se caracteriza, em termos de suporte enquanto um arquivo histórico, constituiu seu acervo a partir de arquivos pessoais de personalidades político-históricas atuantes na história nacional. Nesse artigo analisaremos os impactos promovidos pela digitalização das coleções de fontes primárias da instituição, constando os mais demandados, a potencialização dessa demanda em vista da disponibilidade digital dos mesmos, além da predominância do acesso digital ao total de itens oferecidos a pesquisa.

Atualmente esse acervo se compõe por mais de 1 milhão de documentos disponibilizados à consulta pública, das páginas deste universo cerca de 33% encontram-se digitalizadas e acessíveis online. Considerando o total de páginas o acervo alcança-se cerca de 2.7 milhões de páginas, sendo 920 mil digitalizadas. Sobre este universo que teceremos nossa abordagem e nesse sentido a projeção que a disseminação proporciona ao acervo caracteriza, ao nosso ver, um ativo baseado na ontologia do digital. Dessa forma quando analisamos um recorte de coleções mais consultadas, constatamos que esse atributo – do digital – potencializa simultaneamente a “exponencialida-

de” do acesso, a preservação do suporte original e relativização da custódia desse material, haja vista sua possibilidade de reprodução em massa. Assim recorreremos a um recorte temporal para dedicar nossa observação e, dessa maneira, escolhe-

mos o período de 2010 a 2016. Com efeito, verifica-se que, na totalidade dos cerca de 136 coleções organizadas e disponíveis para pesquisa, que dos 10 fundos mais consultados, 8 já estão disponibilizados para consulta pública no portal CPDOC.

10 ARQUIVOS MAIS CONSULTADOS: 2010 - 2016	
Getúlio Vargas	27,81%
Gustavo Capanema	7,97%
Oswaldo Aranha	7,17%
Anísio Teixeira	5,42%
Ernesto Geisel	4,02%
Alzira Vargas do Amaral Peixoto	3,76%
Antônio Azeredo da Silveira	2,57%
Ernâni do Amaral Peixoto	2,26%
Tancredo Neves	1,99%

Tabela 1 – Arquivos mais consultados

Fonte: DO AUTOR

A tabela acima nos permite observar a demanda predominante recaindo sobre as coleções já com cópias digitais disponíveis. A bem da verdade, a instituição, naturalmente, privilegia a digitalização das coleções mais demandadas, contudo veremos que, mesmo assim, essa qualidade segue ganhando força baseada na disponibilização digital. Assim, em medida semelhante, o movimento de impulsionamento da demanda a partir do oferecimento digital das coleções pode ser sentido em sinal sensível se observado o índice de requisição das coleções quando da comparação em momentos prévios e posteriores às suas respectivas liberações digitais. Dessa forma ilustraremos na sequência esse movimento de potencialização de acesso e demanda.

Como a digitalização e a consecutiva liberação à consulta online se dão por

coleções específicas, nos é possível observar esse movimento em detalhe. No exemplo ilustrado, vemos que o arquivo Ernesto Geisel, devido à relevância histórica de seu titular, sempre desfrutou de grande demanda por seu conteúdo. Mesmo assim a digitalização foi capaz de torná-lo ainda mais acessado, levando-o à segunda posição ante a quarta colocação no ano anterior ao de sua disponibilização online.

Outros exemplos ainda são mais significativos, como no caso da coleção Juarez Távora. Esse acervo ocupava a 29º posição de consulta e, após sua liberação na rede, passou a ocupar a 10º colocação entre os mais consultados. A tabela a seguir exhibe outros exemplos da potencialização concedida pela liberação das cópias digitais na web dos acervos do CPDOC.

ANTES DA DIGITALIZAÇÃO	ARQUIVO	DEPOIS DA DIGITALIZAÇÃO
4º (em 2007)	Ernesto Geisel	2º (em 2008)
29º (em 2008)	Juarez Távora	10º (em 2009)
33º (em 2009)	Paulo Nogueira Batista	19º (em 2012)
16º (em 2009)	Antônio Azeredo da Silveira	8º (em 2012)
28º (em 2013)	Café Filho	11º (em 2014)
20º (em 2013)	Clementini Mariari	15º (em 2014)

Tabela 2 – Impacto proporcionado pela liberação digital das coleções

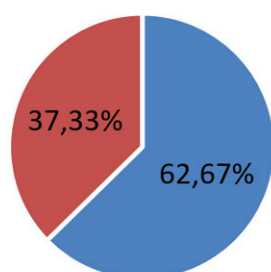
Fonte: DO AUTOR

O aumento da projeção dada a cada coleção específica acaba por projetar a demanda pelo acervo da instituição em sua totalidade. Esse movimento é responsável por ampliar o acesso, diversificar o público e os usos dessas fontes. Mesmo assim, não contraditoriamente, mantêm-se uma política de preservação eficiente por furtar o manuseio dos mesmos. Esse fenômeno complementado pelas novas abordagens criadas para serem atribuídas a esses materiais digitais constituem uma heurística que reconfigura unidades de informação e usuários a partir do suporte. Esse movimento é, para nós, o responsável por criar as bases para o

surgimento de práticas como aquelas discutidas e trabalhadas sob o conceito de Humanidades Digitais^{VIII}.

A amplificação da ressonância caracterizada pelo reaproveitamento das fontes, que podemos constatar no gráfico abaixo, se aplica a usos mais complexos em que elevam a utilização dos agora objetos digitais à patamares automatizados. Esse crescimento de acesso e demanda ajuda a ser explicado também por essa nova forma de uso desses itens, assim, como vemos, a parcela em estado digital (cerca de 33%) corresponde a mais de 60% da demanda.

Total de Consulta Por Suporte



■ Digitalizados ■ Não digitalizados

Gráfico 1 – Volume de consulta por suporte

Fonte: DO AUTOR

Esse processo de demanda ampliada pelo uso dos objetos digitais podem ser compreendido no contexto das Humanidades Digitais, no qual a informação é analisada sob sua representação digital e, através dela, processada informaticamente. Nesse contexto, as Humanidades Digitais se colocaram a refletir essas questões considerando as atividades exercidas no campo da informação conjugadas com as facilidades do uso das tecnologias. Partindo, portanto, das reflexões já trabalhadas sob a perspectiva da Ciência da Informação, mas, agora, com vistas a contribuir com o que se projeta como um novo campo ou como a ontologia do digital, essas atenções agora se voltam especialmente para o universo instrumental do digital compreendido pelos repositórios digitais, curadoria e gestão desses conteúdos, sua preservação e sua disseminação, além de seus impactos nas esperas de ensino e pesquisa acadêmica.

Novas possibilidades de apropriação da informação configuram uma ontologia do digital

A chegada da aplicação de determinadas ferramentas tecnológicas produzindo novos rumos e métodos de pesquisa podem ser considerados o marco mais emblemático do surgimento do rótulo que atende pelo nome de Humanidades Digitais. A soma de recursos oferecidos por softwares com a informação digital proporcionou à pesquisa, ao ensino acadêmico e ao trato informacional produtos inovadores no tocante ao aproveitamento da informação. Nesse sentido temos, por exemplo, as metodologias de georeferenciamento, mineração de dados e textos, a modulação ou rotulagem de tópicos, os processos de visualização da informação entre tantos outros. Todos esses processos ou métodos, quais tem sido aplicados ao consumo da informação em estado digital, compreendem um arcabouço pra-

xiológico e teórico que elevaram as reflexões sobre a presença das tecnologias na realidade da informação a o patamar das Humanidades Digitais.

Vários desses métodos ou processos podem ser aplicados às informações, cada um deles encontra em um determinado repositório informacional uma maior ou menor afinidade no tocante a produção de sentido, além de produtos ou inovação gerados. Para efeito de compreensão, sobre o como as práticas tradicionais das quais a informação tem sido até hoje submetida, contribuíram como alicerces para aplicação das novas técnicas, faremos menção aos métodos de mineração de textos e modelagem de tópicos quais podem ser aplicados, no âmbito ontológico, às informações contidas nas fontes de natureza arquivística.

O método de mineração de textos ou de dados textuais, por exemplo, pode ser compreendido como processo de análise de grandes volumes de texto por software de computador permitindo a extração e visualização dos termos mais recorrentes, possibilitando uma compreensão semântica do documento em curto espaço de tempo. A grosso modo, os softwares de análise de textos são capazes de ler milhares de palavras e informar ao pesquisador quais delas possuem maior recorrência, evidenciando temáticas daquele conteúdo.

Para Moraes e Ambrósio (2007), a mineração de textos é um processo de descoberta de conhecimento, que utiliza técnicas de análise e extração de dados a partir de textos, frases ou apenas palavras. Para os autores esse processo possibilitado pelo desenvolvimento de softwares computacionais, envolve a aplicação de algoritmos que processam textos e identificam informações úteis e implícitas, que normalmente não poderiam ser recuperadas utilizando métodos tradicionais de consulta. (MORAIS; AMBRÓSIO, p.1, 2007)

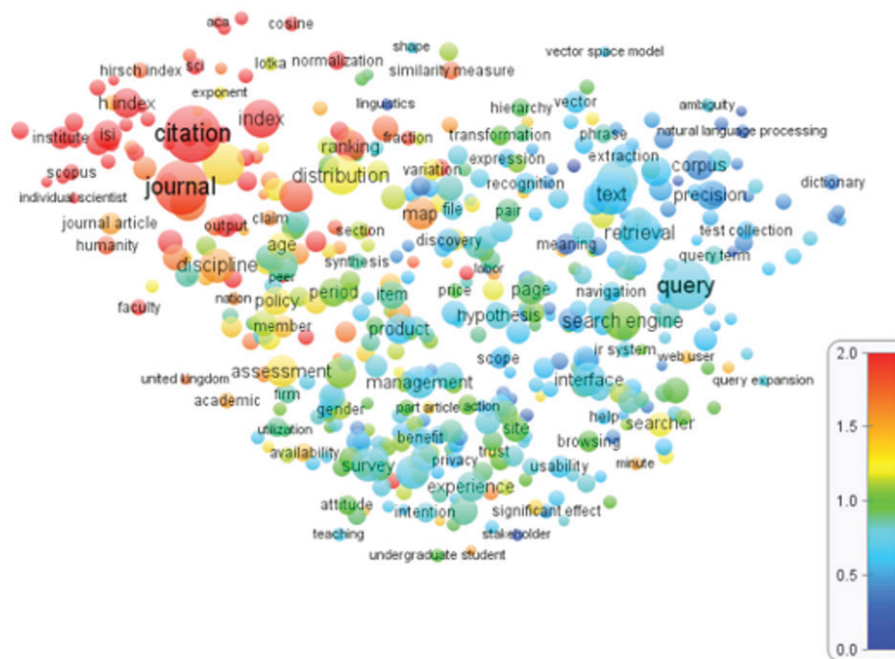


Figura 2 – Exemplo de visualização de resultados textuais minerados com software de computador

Fonte: <http://www.vosviewer.com/text-mining-and-visualization-using-vosviewer>

Uma outra técnica bastante aplicável à produção de novos conhecimentos a partir da análise por máquina é a modelagem de tópicos. A modelagem ou rotulagem de tópicos é o processo realizado por software especializado para descobrir um índice de termos recorrentes para identificação de uma estrutura semântica do documento. O ponto central desse processo reside em agrupar, aqui com o auxílio do pesquisador, uma cadeia de termos minerados sob um identificador comum que o relaciona com os demais. Assim, como bem exemplificam Nolasco e Oliveira (2016), espera-se que termos como “foguetes” e “espaço” estejam associados em maior quantidade ao tópico de “viagens espaciais” do que ao tópico “genética”. Da mesma maneira também os autores sugerem que “genética” contenha mais termos como “gene” e “DNA”. Assim, cada tópico extraído poderia ser inicialmente representado por seus termos mais comuns e poderíamos agrupar documentos de acordo com o respectivo tópico.

Ambas as técnicas podem fazer uso de recursos de visualização como a imagem acima, nos quais é possível analisar um corpus geralmente amplo através da identificação visual de seus termos mais preponderantes. Essas técnicas vêm fazendo parte das práticas de pesquisa acadêmica e projetos especiais de utilização das fontes da instituição. No CPDOC essas práticas já ganham forma e, mesmo ainda numa esfera inicial, vem atendendo projetos específicos de trabalho de extração da informação contida em determinadas partes do acervo da instituição. Além disso, cada vez mais se revela a riqueza do conteúdo dessa massa documental que antes dependeria de muito tempo para ser analisada em grandes escalas. Não obstante a larga digitalização dos conteúdos tem sido condição *sine qua non* para que o desenvolvimento de pesquisas atingisse esse patamar.

Essa transformação ainda se reverbera junto ao perfil dos usuários que cada

vez mais preocupam-se com metadados, OCR e padrões de digitalização. Percebe-se que a preocupação em ler um documento vem migrando para a posse digital do mesmo, importando cada vez mais em tê-lo digitalmente para posteriormente processá-lo em um software.

Atualmente grande parte das pesquisas preocupam-se mais em obter as imagens digitais do que com a consulta através da leitura dos itens. Esse é o cenário pavimentado pelos objetos digitais que foram responsáveis por gerar um arcabouço ontológico que reformatou profundamente unidades de informação, seus usuários e as práticas de consumo dessas fontes bem como as metodologias aplicadas a elas.

Considerações finais

De todo processo de reformatação sofrido pelo ambiente informacional, o de digitalização em massa é o que vem promovendo maiores impactos nas esferas das unidades de informação, dos suportes informacionais e dos perfis de usuários. Por essa razão o processo de preponderância dos objetos digitais deve ser observado e analisado. Das 3 principais esferas alcançadas, o contexto informacionais e apresenta como centro de um processo que gera um novo patamar de abordagem desse objeto. Esse novo patamar é o responsável por transpor a imposição de um finito praticado pelos suportes tradicionais de natureza física. Com a criação de uma massificação do digital, promovendo uma opção clara por esse formato, que promove uma maleabilidade sem precedentes em todos os aspectos de sua constituição, se configura uma transposição a um não-finito que patrocina um novo patamar de abordagem. A esse patamar caracterizamos enquanto ontológico, ao qual se deve inaugurar uma observação sobre seus impactos. Um deles, por exemplo, podemos

vislumbrar nos debates sobre as Humanidades Digitais. Todavia nossa proposta é indicar um construto peculiar na esfera dos objetos digitais que manifestam necessidade de análise e compreensão no âmbito dos estudos informacionais.

Referências

CASTELLS, Manuel. A Sociedade em Rede. São Paulo: Paz e Terra, 2000.

EVENS, Aden. Web 2.0 and the Ontology of the Digital. *Digital Humanities Quarterly*, v.6, n.2, 2012

HUI, Yuk. What is a Digital Object? *MetaPhilosophy: Special Issue: Philoweb: Toward a Philosophy of the Web*. V.43, n.4, Jul, p.380-395, 2012

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

KIRSCHENBAUM, Matthew. The .txtual Condition: Digital Humanities, Born-Digital Archives, and the Future Literary. *Digital Humanities Quarterly*, v.11, n.3, 2013

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

MCCARTY, W. Humanities Computing. In: *Encyclopedia of Library and Information Science*. New York: Marcel Dekker. 2003

MORAIS, E. A.; AMBRÓSIO, A. P. L. *Mineração de Textos*. Relatório Técnico. Instituto de Informática; Universidade Federal de Goiás, 2007. Disponível em: <http://www.inf.ufg.br/sites/default/files/uploads/relatorios-tecnicos/RT-INF_005-07.pdf>. Acesso em: 02 jun. 2017.

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

NOLASCO, Diogo; OLIVEIRA, Jonice. Modelagem de tópicos e criação de rótulos: identificando temas em dados semi-estruturados e não-estruturados. In: OGASSAWARE, Eduardo; VIEIRA, Vaninha (orgs.). *Tópicos em gerenciamento de dados e informações*. Salvador: Sociedade Brasileira de Computação, 2016. cap. 4, p. 87-112. E-Book. Disponível em: <<http://sbbd2016.fpc.ufba.br/e-book/minicursos.pdf>>. Acesso em: 02 jun. 2017.

ROSS, Lyman; SENNYEY, Pongracz. The Library is Dead, Long Live the Library! The Practice of Academic Librarianship and the Digital Revolution. *The Journal of Academic Librarianship*, v. 34, n.2, p. 145–152.

THIBODEAU, K. *Overview of Technological Approaches to Digital Preservation and Challenges*. Coming Years The State of Digital Preservation: An International Perspective. Anais... Washington: CLIR and Library of Congress, 2002.

VARELA, Aida Varela. *A explosão informacional e a mediação na construção do conhecimento*. Encontro Nacional de Pesquisa em Ciência da Informação (ENANCIB), 6., 2005, Florianópolis, SC. Disponível em <<http://enancib.ibict.br/index.php/enancib/vienancib/paper/viewFile/1755/896>> Acesso em 7 set 2018.

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

Recebido em 19/12/2018
Aprovado em 24/01/2019

I Renan Marinho de Castro. Mestre em História, Política e Bens Culturais pela Fundação Getúlio Vargas (FGV). Coordenador do Programa de Arquivos Pessoais da FGV, Rio de Janeiro, Brasil. Contato: renan.castro@fgv.br

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

III O termo *Born-Digital* (em tradução literal 'Nascido Digital') em sua conceituação original em língua inglesa refere-se ao suporte produzido em meio digital e apenas através deste acessível, por essa razão elegemos nos referir neste texto ao termo original em língua inglesa '*Born Digital*'.

IV Podemos considerar enquanto um intercâmbio de "bens comuns" de natureza metodológica que aproximou campos de pesquisa em torno da descoberta de novos métodos informáticos, motivando uma configuração de conexão entre campos disciplinares(MCCARTY, 2003, p.1).

V Nos apropriamos do conceito sob o contexto filosófico para tratar não "do *ser enquanto ser*" "concebido em uma natureza comum", mas de um construto tecnológico que atende pelo nome de objeto digital que, ao reformular o papel da informação, se apresenta passível do estabelecimento de uma nova heurística ao qual ainda estamos nos dando conta e que precisamos compreender em sua essência.

VI Cf. Svensson, 2009, p.10-11.

VII O Centro de Pesquisa e Documentação de História Contemporânea do Brasil (CPDOC) da Fundação Getúlio Vargas foi criado em 1973 com o objetivo de abrigar conjuntos documentais relevantes para a história recente do país e desenvolver pesquisas históricas, tendo inicialmente seu próprio acervo como fonte privilegiada de consulta. Os conjuntos documentais doados ao CPDOC, que podem ser conhecidos na íntegra através da internet, constituem, atualmente, o mais importante acervo de arquivos pessoais de homens públicos do país, reunindo milhares de documentos. A organização desses arquivos e sua abertura à consulta pública, hoje totalmente informatizada por meio do sistema Accessus, são tarefas primordiais do Centro

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