Emanoel Cesar Pires de Assis Cláudio Augusto Carvalho Moura Isabela Melim Borges Sandoval (ORGANIZADORES)

# HUMANIDADES DIGITAIS

Leitura e Tecnologia



Emanoel Cesar Pires de Assis Cláudio Augusto Carvalho Moura Isabela Melim Borges Sandoval (ORGANIZADORES)

# 

Leitura e Tecnologia

1ª Edição Tubarão/SC 2014



#### Humanidades digitais: leitura e tecnologia 1ª edição - 2014

### © Copyright 2014 by Emanoel Cesar Pires de Assis, Cláudio Augusto Carvalho Moura e Isabela Melim Borges Sandoval

#### **Editores**

Emanoel Cesar Pires de Assis Cláudio Augusto Carvalho Moura Isabela Melim Borges Sandoval

#### Conselho Editorial

#### Presidente

Carlos Alberto Antunes Maciel (Université de Nantes)

#### Membros

Maria Eunice Moreira (PUCRS) Terezinha Maria Scher Pereira (UFJF) Zahidé Lupinacci Muzart (UFSC) Cristiano de Sales (Mackenzie) Marly Gondim Cavalcanti Souza (UESPI)

Maria Aparécida Donato de Matos (UFRJ) Francisco Wellington Borges Gomes

Luiz Guaracy Gasparelli Junior (FAFIMA)

Tânia Regina Oliveira Ramos (UFSC)

Projeto gráfico e capa Otávio Guimarães Tavares

Diagramação Mairla Pereira Pires Costa Revisão Dos editores

Impressão e acabamento Gráfica e Editora Copiart

H98 Assis, Emanuel Cesar Pires de.

> Humanidades digitais: leitura e tecnologia / Emanoel Cesar Pires de Assis; Cláudio Augusto Carvalho Moura; Isabela Melim Borges Sandoval (organizadores). - Tubarão: Ed. Copiart; Florianópolis : NuPILL/UFSC, 2014.

244p.: il.; 21 cm. ISBN: 978.85.8388.022.6

1. Literatura e tecnologia. 2. Literatura e Internet. 3. Literatura digital. 4. Bibliotecas digitais. 5. Repositórios institucionais. I. Assis, Emanoel Cesar Pires de. II. Moura, Cláudio Augusto Carvalho. III. Sandoval, Isabela Melim Borges.

CDD (22. ed.) 801.959

Elaborada por Sibele Meneghel Bittencourt - CRB 14/244

Todos os direitos reservados e protegidos pela Lei 9.610 de 19/02/1998. É proibida a reprodução parcial ou integral desta obra, por quaisquer meios de difusão, inclusive pela internet, sem prévia autorização dos editores.

# **SUMÁRIO**

# PARTE I – LITERATURA, LEITURA E TECNOLOGIA

Leitura e tecnologia <eros &="" psique="">: leituras jovens do mundo13 Miguel Rettenmaier e Tania Rösing</eros>
A narrativa transmídia como estratégia de incentivo à leitura36 Elisα Cristinα Delfini Corrêα
Sobre a paratextualidade: um relato de teoria aplicada em sala de aula
54 Cláudio Augusto Carvalho Moura
Leitura digital em tela: a hipermodalidade como forma de significação textual
PARTE II - FERRAMENTAS E REPOSITÓRIOS DIGITAIS
O uso de elementos semânticos no processo de recuperação da informação em ambientes digitais
Bibliotecas e repositórios digitais: reflexões, tecnologias e aplicações 130 Divino Ignacio Ribeiro Jr.
Adaptação de conteúdo para Dispositivos Móveis

# PARTE III – ESTATÍSTICA TEXTUAL

Metodologia, ciências duras e literatura em meio digital Saulo Cunha de Serpa Brandão	197
Métricas rosianas (1929-1967) Verônica Ribas Cúrcio	213
GOOFRE versão 2: ver e tratar 600 bilhões de palavras Etienne Brunet, Laurent Vanni, Cláudia Grijó Vilarouca (Traducão)	222

# **APRESENTAÇÃO**

Este volume é fruto dos trabalhos do Núcleo de Pesquisa em Informática Literatura e Linguística da Universidade Federal de Santa Catarina (NuPILL/UFSC), seus parceiros e convidados, além dos participantes do II Congresso Internacional e VI Congresso Nacional de Literatura e Informática, realizado pelo núcleo em dezembro de 2013 nas dependências da UFSC. Sua organização e publicação contaram com o apoio da Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES), que, junto a outros parceiros, tornou também possível a bem-sucedida realização do nosso simpósio. São esses: a Fundação de Amparo à Pesquisa e Inovação do Estado de Santa Catarina (FAPESC), Pró-Reitoria de Pós-Graduação (PRPG/UFSC), Programa de Pós-Graduação em Estudos da Tradução (PPGET/UFSC), Programa de Pós-Graduação em Literatura (PPGL/UFSC), a direção do Centro de Comunicação e Expressão (CCE/UFSC), Departamento de Língua e Literatura Vernáculas (DLLV/UFSC), Departamento de Língua e Literatura Estrangeiras (DLLE/UFSC) e o Núcleo de Literatura e Memória (NuLIME/UFSC).

Para compor o volume em questão foram selecionados alguns textos provenientes dos eixos *Leituras em meio digital e Estatística textual*, aqui agrupados em conjuntos distintos, conforme as proximidades temáticas de cada texto. Apresentamos, ao todo, 10 capítulos, divididos em 3 seções: *Literatura. Leitura e Tecnologia, Ferramentas e Repositórios Digitais e Estatística Textual*, com textos de docentes e pesquisadores de variadas instituições que possuem, como elo comum, o interesse na pesquisa voltada à tecnologia e suas implicações seja na literatura, seja em questões correlatas.

No primeiro capítulo da sessão inicial, Miguel Rettenmaier e Tania Rösing tomam como pano de fundo as Jornadas Literárias de Passo Fundo para discutir as relações dos jovens com a tecnologia por meio da literatura, e vice-versa, discussão que se soma à análise da obra *Eros & Psiquê.com.br* (2011). Na sequência, Elisa Cristina Delfini Corrêa faz um percurso da contação de histórias tradicional à narrativa transmídia, enfatizando seu potencial como estratégia para incentivo de leitura. O capítulo seguinte, de Cláudio Augusto Carvalho Moura, relata um experimento teórico-prático sobre a paratextualidade, realizado em sala de aula com o uso de minicontos disponíveis online. Fechando a sessão, Gisella Meneguelli Sousa parte de uma experiência com Educação a Distância (EaD) para refletir sobre o letramento digital, e apontar a produtividade do *interactive model* no processamento da leitura de textos nos quais são utilizados recursos hipermodais.

A sessão seguinte inicia com o capítulo de José Eduardo Santarém Segundo acerca do uso de elementos semânticos no processo de armazenamento e recuperação da informação em repositórios digitais através do processo de Folksonomia Assistida e do modelo Representação Iterativa. Em seguida, Divino Ignacio Ribeiro Jr.

apresenta a evolução da discussão sobre os repositórios e bibliotecas digitais, suas particularidades, discutindo questões ligadas à importância de sua estruturação, manutenção, implementação e gerenciamento eficientes. O último capítulo da sessão, de autoria de Fábio Bif Goularte, Beatriz Wilges e Silvia Modesto Nassar, apresenta um protótipo de caderno de material didático adaptado para dispositivos móveis ainda em fase de testes.

Na última sessão deste volume temos o capítulo inicial de Saulo Cunha de Serpa Brandão, que discute a necessidade de uma delimitação metodológica eficaz para os estudos literários e o uso de ferramentas telemáticas para análise de literatura em meio digital. O capítulo seguinte traz um estudo de estatística textual de Verônica Ribas Cúrcio sobre a obra de Guimarães Rosa, apontando características do seu vocabulário e comparando os resultados encontrados com três teses da crítica literária rosiana. O capítulo que encerra tanto a sessão quanto o livro tem como autores Etienne Brunet e Laurent Vanni. Nesse capítulo é relatada a empreitada, em curso, de análise e tratamento de *corpora* linguísticos disponíveis *online*, tomando como partida a língua francesa.

Embora não possamos chamar de recente toda a revolução tecnológica que presenciamos, vez que perscrutadas, suas raízes indicam, pelo menos, dois séculos de profundidade; tampouco são recentes os estudos que tratam do tema. Mesmo se falarmos de informática, teremos já meio século, e até mesmo a *internet* em si já não pode, ou pelo menos não deveria, ser tratada como novidade, visto que o início de sua massificação já tem em torno de duas décadas. Contudo, é na constante evolução desses objetos que encontramos um campo fértil, repleto de questionamentos a serem respondidos, ou, pelo menos, caminhos a serem apontados, na busca por algumas respostas; quando

não, questões. Neste sentido, este livro tem por objetivo contribuir com a ampliação das discussões acerca dos pontos de entrecruzamento da tecnologia com nossas áreas de interesse.

Os organizadores.

# PARTE I LITERATURA, LEITURA E TECNOLOGIA

# Leitura e tecnologia <*Eros & Psique*>: leituras jovens do mundo

Miguel Rettenmaier<sup>1</sup>
Tania Rösing<sup>2</sup>

### As Jornadas Literárias de Passo Fundo: das ruas ao rolezinho

De todas as Jornadas Literárias de Passo Fundo, nos seus 32 anos de história, a edição de 2013, com o tema "Leituras jovens do Mundo", talvez tenha sido a que mais intimamente se associou com o contexto, dialogando, no calor do momento, com acontecimentos transcorridos tantos dias antes, quanto pouco depois do encontro. Em junho de 2013, poucas semanas antes da Jornada, manifestações populares despontavam nas grandes metrópoles brasileiras, irradiando-se por todo o país. Tais manifestações, sem partido político definido, sem causas específicas, sem linha ideológica reconhecível, fragmentadas e heterogêneas, levaram às ruas milhões de pessoas, em grande parte jovens, mobilizados por um sentimento de insatisfação quanto aos rumos dos enfrentamentos a problemas brasileiros históricos. Os protestos, relacionados a uma grande variedade de temas, envolviam a

<sup>1</sup> 

<sup>&</sup>lt;sup>1</sup> Professor do Programa de Pós-Graduação em Letras da Universidade de Passo Fundo (UPF), coordenador do Acervo Literário de Josué Guimarães (ALJOG/UPF), membro da Comissão Organizadora das Jornadas de Literatura de Passo Fundo.

CONTATO: mrettenmaier@hotmail.com

<sup>&</sup>lt;sup>2</sup> Professora do Programa de Pós-Graduação em Letras da Universidade de Passo Fundo (UPF), coordenadora do Centro de Referência de Literatura e Multimeios - Mundo da Leitura da UFP, idealizadora e coordenadora das Jornadas Literárias de Passo Fundo. Bolsista de Produtividade em Pesquisa do CNPq. CONTATO: tmkrosing@gmail.com

precariedade e o preço do transporte público, os gastos com grandes eventos esportivos internacionais, como a Copa do Mundo de 2014, a má qualidade dos serviços públicos e a corrupção em todas as esferas de poder. Fomentadas pela "propagação viral" potencializada pelas redes sociais, as manifestações de junho de 2013 foram assunto na imprensa internacional, obrigando a reorientações nos governos em nível federal, estadual e municipal.

Também por força das redes sociais, no início de 2014, um tipo diferente de manifestação, envolvendo também os jovens, passou a acontecer em outro espaço de convívio social: o shopping center. No foi conhecido como "rolezinho", milhares de jovens, majoritariamente advindos da periferia, por meio das redes sociais, marcavam encontros nos grandes shopping centers do país, congestionando desordenadamente os espaços destinados aos consumidores. Sem as conotações políticas de junho de 2013, por mais amplas que fossem no que se colocava como pauta de reivindicações ou como inventário de descontentamentos, o rolezinho teria como origem a busca de diversão por jovens afastados dos grandes centros econômicos e não completamente integrados às possibilidades do consumo. Diminutivo de "rolé", pequeno passeio, na linguagem informal de determinados grupos jovens, os encontros foram frontalmente questionados por parte da opinião pública e da imprensa brasileira, que associou as aglomerações, ou *flash mobs*, à delinguência e à criminalidade. Para alguns formadores de opinião, os encontros eram e são, como declarou Rodrigo Constantino, obra de uma "barbárie" protegida pelo "manto do preconceito":

Uma turba de bárbaros invadindo uma propriedade privada para fazer baderna não é protesto ou "rolezinho", mas invasão, arrastão, delinquência. O primeiro passo para vencer esse avanço da barbárie é chamá-la pelo nome certo.

Selvagens que cospem na civilização não são "manifestantes" coisa alguma.<sup>3</sup>

Antes de se resolver a questão que se impõe, de natureza tão profunda, na forma de uma posição segundo a qual os sujeitos que integram o rolezinho sejam tipos que "não toleram as 'patricinhas' e os 'mauricinhos', a riqueza alheia, a civilização mais educada", deve-se observar dois pontos, como elementos de uma discussão sempre limitada no que de provisório há em qualquer conclusão quando se envolvem os jovens e a contemporaneidade. Em primeiro lugar, deve-se pensar no shopping center como um espaço em que há mais do que o mero consumo em um espaço que "tem dono". Ali, na disposição arquitetônica, no design das fachadas, dos logotipos, das marcas, nos produtos oferecidos e cultuados, no trânsito incessante das pessoas, há um jogo de produção de valores envolvido em uma nova significação de consumo. O shopping é lugar de uma antropologia a ser discutida. Isso o torna um reduto muito mais complexo do que mero "local agradável para o consumo e o lazer das famílias e indivíduos". Massimo Canevacci, antropólogo italiano, um dos convidados da 15ª Jornada Nacional de Literatura de Passo Fundo, ao circular em 1984, em São Paulo, no shopping Eldorado, percebeu, "naquela espécie de astronave sem janelas aterrissada num cruzamento estratégico" (CANEVACCI, 2013, p. 259), uma promessa e um sentido novo ao consumo, que não mais podia ser visto segundo a postura econômica clássica, como a fase final do ciclo iniciado pela produção:

As experiências em uma área como o Eldorado eram outras, claro, também de consumo, mas principalmente se comunicava uma espécie de praça metropolitana onde as distinções clássicas de público-privado não tinham mais sentido. É onde os consumidores eram ao mesmo tempo

<sup>3</sup> "O 'rolezinho' da inveja. Ou: A barbárie se protege sob o manto do preconceito". Disponível em http://veja.abril.com.br/blog/rodrigo-constantino/cultura/o-rolezinho-da-inveja-ou-a-barbarie-se-protege-sob-o-manto-do-preconceito.

O texto, deliberadamente, não se encontra nas referências ao final deste artigo.

performers que atraíam um ao outro reciprocamente. Consumiam-se olhares recíprocos (CANEVACCI, 2013, p. 259).

Para Canevacci, era necessário pensar o shopping com a mesma seriedade que Marx pensara a produção, no sentido de se rediscutirem as definições tradicionais de classes sociais, as relações dialéticas entre público e privado, estrutura e superestrutura, Estado e sociedade civil, religião e política, direito e cidadania: "No cotidiano, o Eldorado era mais forte do que a fábrica. O cinturão operário do ABC, as três cidades que circundam a Grande São Paulo, cheias de fábricas, não era mais o centro da política" (2013, p. 259).

Em segundo lugar, quando o centro se desloca do mundo adulto, que corresponderia ao trabalho, para um campo onde circulam jovens, a questão deve buscar entender o que é ou o que são esses sujeitos, quais formas artísticas a eles se associam. Isso implica pensar com e sobre alquém para o qual já se delegou o futuro, quando hoje é cada vez mais seu o presente. A temática da 15ª Jornada Nacional de Literatura, "Leituras Jovens do Mundo", a partir das perspectivas que cercavam o sujeito jovem, exigiu de si livrar-se dos estereótipos em torno da palavra, buscando senão conceitualmente, pela impossibilidade de uma pela heterogeneidade da realidade, uma linha definição entendimento que tratasse do que se poderia vislumbrar como "jovem". As relações abertas a partir dessa preocupação trouxeram elementos instigantes no que superava afirmativas generalistas sobre a juventude, sobre arte e sobre literatura. Principalmente pelas condutas contemporâneas, intimamente associadas à tecnologia, ser jovem não estaria mais anexado irrevogavelmente a uma faixa etária definitiva. E isso tinha implicações na arte e na obra literária.



Figura 1. Folder 15ª Jornada Nacional de Literatura de Passo Fundo. Fonte: Zero3 Comunicação e Design

# Ser jovem, extremo e sincrético4.

As Jornadas de Passo Fundo, em seu histórico, no trabalho de promoção da leitura e da literatura, nunca se eximiram de defender uma posição pela qual se defendia indissociabilidade entre tecnologia, sociedade e cultura. A tecnologia mobilizara, em 2001, a temática da  $g^a$  Jornada Nacional de Literatura, que versava sobre a galáxia de Gutenberg, invadida por dois impactantes elementos: um novo sujeito e os novos suportes digitais de leitura, ainda em seus primeiros passos no Brasil. Em 2009, a temática vinculando o livro impresso e a robótica foi

\_

<sup>&</sup>lt;sup>4</sup> Parte da base teórica deste trabalho foi apresentada em comunicação no 5ª Simpósio Hipertexto e Tecnologias na Educação, realizada no campus da Universidade Federal de Pernambuco (UFPE) em novembro de 2013.

"Arte e tecnologia: novas interfaces". Em 2011, com a temática "Leitura entre nós: redes, linguagens e mídias", a palavra "nós" aludia tanto à conexão da informática globalizada quanto ao indicativo de toda uma comunidade voltada à formação dos leitores nos 30 anos de história das Jornadas de Passo Fundo. A consciência da importância das tecnologias tinha forte consistência teórica, vinculada a um pensamento, de certa forma, visionário. Por Lévy, confirmava-se a cada período de tempo, a cada transformação, a cada inovação, a noção de que as tecnologias são produtos de uma sociedade e de uma cultura, da mesma forma como seria apenas conceitual a distinção entre a cultura (a dinâmica das representações), a sociedade (as pessoas, seus laços, suas trocas, suas relações de força) e técnica (artefatos eficazes) (1999, p. 22). Mais do que no plano conceitual a associação de tais termos repercutiria virtualmente na política numa referência na qual as TICs poderiam ser orientadas para a construção de uma democracia em tempo real, diferente da atual democracia representativa. Nas novas dinâmicas, a potência suplantaria o poder, a espécie humana constituir-se-ia em hipercórtex, pelo qual no ciberespaço aprofundar-se-ia a democracia quando a cada cidadão seria legado o direito de

contribuir de maneira contínua para a elaboração e o aperfeiçoamento dos problemas comuns, para a abertura de novas questões, para a formulação de argumentos, para enunciar e adotar posições independentes uma das outras sobre uma variedade de temas (LEVY, 2007, p. 65).

Esse projeto, contudo, paradoxalmente, como utopia, traria em si a instabilidade dos movimentos coletivos, o que não lhe garantiria o sucesso de uma constituição a-histórica, no qual se superariam todos os conflitos pela determinação final do bem comum. A cibercultura é sempre uma solução parcial, em si mesma, campo de problemas e conflitos. Nesse sentido, por si, as TIC não teriam o poder ou a potência de conter o poder. O progresso industrial, justamente por estabelecerse como eixo do fluxo da economia capitalista, seguidamente foi avaliado como legitimador de uma determinada ordem social, a partir

da qual os próprios artefatos, mais do que ferramentas de produtividade e eficiência, representariam uma semântica legitimadora do poder constituído. Para Feenberg, a tecnologia, como manifestação de uma racionalidade política, como elemento ambivalente, pode tanto servir ao controle e à hegemonia, quanto à transformação cultural. Hegemonia e transformação estariam associadas, respectivamente, à autonomia operacional (operational autonomy), que serve como base estrutural de dominação, e às margens de manobra (margin of maneuver), tecnicamente mediadas para redefinir formas, ritmos e propósitos. No que se refere às margens de manobra, a tecnologia seria operada para transformação civilizadora da sociedade apenas e só mediante o domínio autoconsciente dos sujeitos quanto à técnica, mesmo porque, como ambivalente, estaria vinculada a um código social. Como a própria linguagem, os códigos técnicos do capitalismo não têm poder absoluto ou intrínseco, e sua hegemonia é sempre relativa, dada a ambiguidade de seus fundamentos, de suas estruturas e de suas manifestações. Assim, a arte literária, articulada à tecnologia e a seus códigos, pode conduzir a que se manipulem conscientemente estratégias, retirando das tecnologias a aura fáustica que ainda podem possuir, para contribuir para uma transformação cultural civilizatória.

A abordagem de Feenberg permite, nesse sentido, que se vislumbre o jovem como sujeito e agente transformador, vinculado a uma condição estética de renovação de valores. Segundo Canevacci (2005), a sociedade de hoje exige uma dilatação do conceito de jovem, em nome de uma acepção que, no cenário de uma metrópole comunicativa e imaterial, seja plural, diferenciada e móvel. O jovem líquido da contemporaneidade deflui por interzonas, é um "eXtremo interminável", um processo sincrético associado a múltiplas influências em movimento indefinido e imprevisto. À noção de pessoa "multivídua", correspondem esses trânsitos de vida que desconhecem ou não obedecem conceitos. Para o pensador italiano, a base dos conceitos e das divisões é política e esconde relações de um poder, hoje, fragilizado ou fracionado.

O ponto de vista dialético da modernidade contrapunha polos e apresentava sínteses em um fluxo linear e natural da história e do tempo, no plano íntegro de um panorama visível ou, ao menos, perceptível. Mas tudo isso estaria hoje definitivamente problematizado. No universo dos fragmentos líquidos, das culturas intermináveis, as entidades eX-terminadas inviabilizam toda série de conceitos e de definições que unificam e tranquilizam a perspectiva a uma amostra constatável e delimitada de fenômenos. O controle e a identidade fixa padecem na metrópole e na mídia. A localização única danifica-se, as faixas e distinções fragilizam-se, as separações etárias, de gênero, de proveniência, de trabalho, todas elas, é importante que se diga, essencialmente políticas, estão agora confusas. As classes sociais não se explicam mais. O corpo é um aparato de signos, continuação das paredes dos quartos juvenis e dos muros pichados, o piercing e a tatuagem ultrapassam e contaminam o orgânico – o corpo é prótese, a prótese é corpo. Sendo corpo e prótese o mesmo fracionado e o outro introduzido, o particular, o individual, é tatuado e perfurado pelo externo, pelo social. Mas esse social não é político. As culturas juvenis estão abertas demais para obedecerem a referências precisas e a determinações diretivas. Sua forma de viver está nos "fluxos desordenados e polifônicos/dissonantes da comunicação metropolitana e dos panoramas midiáticos" (CANEVACCI, 2005, p. 27), como pode se perceber nas movimentações populares de junho de 2013, contemporâneas à temática da Jornada.

De forma semelhante, no plano da linguagem e da arte, não há síntese, mas opostos associados que não dizem a que pertencem, não há determinada manifestação artística. As coisas se somam e somando-se, se dividem, se separam, se singularizam. Os códigos estão remixados, nenhuma dominante detém exclusivas propriedades, não há mais música sem imagem, imagem sem música, palavra sem arquitetura, pintura sem indústria, arte sem instalação. A folha é eletrônica, a tela do computador é feita de páginas, mas rola como papiro. Vide o *e-space*:

No e-space posso entrar em contato com qualquer um sem conhecê-lo e ir a qualquer lugar sem ir. O qualquer-lugar, a ubiquidade se insere no e-space. As fronteiras furam. Posso combinar infinitos textos escolhidos ou encontrados casualmente que acho em minha navegação, que nunca teria chegado a conhecer em técnicas anteriores. O copyright como monopólio rígido de instituições editorias entra em crise com as novas tecnologias: posso citar tudo e todos sem ter de me submeter às compras tradicionais. Citando as fontes, o que é uma atitude correta inestimável, posso juntar textos antes inimagináveis. Agora os encontros no e-space. (CANEVACCI, 2005, p. 167-68).

Nesse novo "pluriverso", as práticas da cultura digital transformam a cidade industrial em metrópole flutuante e as identidades fixas em sujeitos diaspóricos:

A transformação da cidade industrial em direção a uma metrópole flutuante é o fieldwork em que se deve colocar o trabalho etnográfico. As práticas da cultura digital nos sites da web; o fim do trabalho fixo e a afirmação de identidades fluidas; as subjetividades performativas no consumo e nas estéticas; as individualidades diaspóricas que atravessam e cruzam lugares, espaços, zonas, interstícios; os processos de hibridização entre polifonias e dissonâncias: tudo delineia a afirmação de uma metrópole comunicacional (CANEVACCI, 2013, p. 271-72).

Nesse ambiente de culturas "glocais", articuladas entre o global e o local, o sujeito vaga, contamina-se, libera-se de matrizes e provoca a evaporação de tudo que é conhecido e reconhecido, as referências teóricas, políticas, afetivas, de amizade — a esfera do que é já experimentado e do autocontrole começa a se descosturar pelos imprevistos percursos de um eu incerto que aceita perder os próprios sistemas de equilíbrio. Seu trânsito, contudo, não é o do migrante caracterizado pela penúria e pela necessidade de, em outro local, folclorizar a si mesmo na defesa, às vezes desesperada, da própria

identidade. Pelo contrário, o sujeito vagante "exprime uma tendência de uma minoria não minoritária que sai de seu mundo local, que quer conhecer e atravessar as diferentes culturas e, ao mesmo tempo, a si próprio" (CANEVACCI, 2013, p. 114). Nessa cidadania transitiva de automontagem, pelo poder da comunicação digital, o eu que coloca em crise as tendências à verticalização expande a "liberdade sincrética", capaz de realizar mesmo uma ortografia em síncrese:

SincrétiKa, assim, manifesta o seu "K" luminoso que inflama e um "a" plural que expande a disparidade dos "objetos" encontrados ou descobertos nos itinerários transfigurados. SincrétiKa exprime o vagar do etnógrafo e o vago da arte. SincrétiKa redefine a política atual caracterizada pela comunicação digital: as relações entre auto-representação e comunicação é política, é política comunicacional cada vem mais experimentada e difundida graças às tecno-culturas da montagem (CANEVACCI, 2013, p. 14 – Grifos do autor).

A web, a internet, a ausência de limites nos percursos hipertextuais, a falta de costura precisa entre as coisas e os sujeitos, as fronteiras perturbadas entre os gêneros, entre ficcional e real, entre o ensaio e a poesia, entre a prosa artística e a fala das ruas, a referência initerruptamente parodiada, as contaminações de toda a ordem, tudo isso colocou a assinatura e a autoria em crise, a ponto de se imaginar que a arte teria como futuro a produção coletiva, a coautoria grupal, em um tipo de poiéses conectada ao máximo limite de volatilização do individual. Entretanto, se "contexto midiático é parte da experiência do sujeito" (CANEVACCI, 2013, p. 173), se é certo que ninguém vive sem mídia ou fora dela, é certo também que as tecnoculturas afirmam a subjetividade de cada navegante. Em outros termos: há um elemento singular-plural nas subjetividades multividiais. A sincrética do interior do eu ecoa, dissonante, na arte vagante.

O cenário que está emergindo, cruzando arte e digital, oferece diferentes perspectivas inéditas. Uma geração de artistas, designers, performers, arquitetos, sond-designers (cujas experimentações disciplinares são desafiadas e atravessadas) começa a plasmar obras sem "refletir" sobre a mídia generalista, cada vez menos hegemônica justamente entre as jovens gerações. (CANEVACCI, 2013, p. 175).

As produções sincréticas, polifônicas e diaspóricas reproduzem-se sem perder a força "aurática". Nessa circunstância, segundo o teórico, **em vez de arte coletiva, se afirmam "artistas conectivos"** (CANEVACCI, 2013, p. 176). É talvez o caso de Guto Lins, escritor e designer também presente em agosto de 2013, em Passo Fundo, na 15ª Jornada Nacional de Literatura. É de sua autoria a obra *Eros e Psiquê.com.br*, publicada pela Manifesto Design, em 2011, recontando uma antiga história, sincreticamente.

### Do amor e da alma: mito, livro e tela.

Há o mito e tantas possíveis lições quanto um mito pode permitir: Eros e Psiquê. A história conta que uma princesa, a mais jovem de três irmãs, permanecia solteira, mesmo sendo tão bela quanto "a linguagem humana se mostrava incapaz de descrever" (APULEIO, 1956, p. 23). Digna, pela beleza, de honrarias, como uma imortal, Psiquê afronta, mesmo sem querer, Vênus, que pede ao filho que a vingue, apaixonando a jovem pelo último dos homens. O poder do amor, contudo, não tem a possibilidade de resguardar-se de si mesmo. Deixada em um rochedo por determinação do oráculo do Templo de Mileto, para ser possuída por um monstro horrível "fero como a serpe" (APULEIO, 1956, p. 28), a princesa é salva pelo vento e conduzida ao palácio do Cupido, do qual se torna esposa, mesmo sem a possibilidade de vê-lo.

Solitária em um palácio em que lhe servem vozes sem corpo, Psiquê recebe a frequente visita do marido sem jamais saber quem é. Logo, fica sabendo por Eros que terá filho que poderá, por seus méritos, como mulher, ser tanto um deus, como um mortal. É, então, avisada pelo marido sobre os perigos da inveja, agora das próprias irmãs, que

desconfiam ser um deus o esposo da mais jovem. Mesmo assim, Psiquê se deixa enganar. Em uma das visitas, é convencida a ver e a imolar o marido, com uma lâmpada e um punhal, segundo as irmãs, uma "serpente enroscada em inúmeros anéis, com a fauce cheia de sangue venoso e a boca larga como um abismo" (APULEIO, 1956, p. 46), forma outra da mesma predição pítica<sup>5</sup> que já a condenara.

Aproveitando o sono do marido, Psiquê ilumina seu corpo e descobre ali, não um ser monstruoso, mas um "liso e luminoso" corpo de deus. Embevecida pela beleza do marido, Psiquê fere-se nas suas armas, novamente desvirginando-se, agora pelo sangramento em uma das setas; da mesma forma, involuntariamente queima o corpo do filho de Vênus, que desperto, alarmado e ferido pelo óleo fervente, parte, abandonando a princesa no exterior do castelo. Persuadida por Pã a não se abandonar à morte, Psiquê vai às cidades das irmãs e mente, a cada uma, terem sido objeto do amor do deus. Esperando serem carregadas por Zéfiro ao palácio do Amor, as irmãs se atiram para morte em um abismo.

Psiquê, neste momento, passa a percorrer o mundo em busca do marido, enquanto Vênus, alertada por uma gaivota, no mar, descobre o filho enfermo e o censura fortemente. Frente sua cólera, Ceres e Juno tentam abrandar seus sentimentos alegando que amor e a juventude são coisas que não podem ser reprimidas:

Quem fez seu filho, senhora, para que tanto insistas em contrariar-lhe os prazeres a ponto de quereres perder aquela a quem ele ama? Será porventura crime ter sorrido a uma linda menina? Ou ignoras que ele já é um adolescente e até um adulto? Esqueceste a sua idade? Parece-te um menino por carregar tão lepidamente os seus anos? (APULEIO, 1956, p. 59)

<sup>&</sup>lt;sup>5</sup> Pítica, de Pítio, um dos sobrenomes de Apolo.

A intervenção das deusas não auxilia Psiquê que, capturada, deve vencer provas, sendo também molestada pelas servas de Vênus, Inquietação e Tristeza. Precisa, então, separar sementes misturadas em enorme porção, extrair um floco de lã de ouro de ferozes ovelhas, trazer à deusa uma ânfora cheia de água dos rios infernais e, finalmente, descer aos infernos, trazendo uma caixa fechada com parte da beleza de Prosérpina. Psiquê realiza todos os trabalhos, ajudada ou por elementos da natureza ou por emissários divinos, mas termina por trair a si mesma. Retornando do Tártaro por um respiradouro, não se contém e abre a caixa, contrariando ao que lhe fora determinado. É envolvida pelo sono infernal, sendo salva pelo Amor até o perdão de ambos, princesa e deus, por Júpiter, que ainda concede a ela a imortalidade e ao casal uma união perpétua. Desse casamento, nasce uma filha, chamada Volúpia.

O mito do Amor, Eros, e de Psiquê, redigido por Apuleio, tem sido lido como a união entre alma e amor como intenção simbólica sob a fabulação. Há, obviamente, elementos que se relacionam a elementos do folclore universal, envolvidos nos conflitos entre mortais e deidades, ou entre irmãos por poder e riqueza, nos tabus e proibições sexuais, entre eles a virgindade e a nudez do homem frente à mulher, nas provas a que se submetem os heróis, nas contravenções cometidas pelos protagonistas, nas imposições do destino ao filho mais moço. Amor e Psiguê, contudo, acrescentam um elemento importante: o casal de apaixonados é ainda jovem. O Amor é, segundo Júpiter, um adolescente a quem lhe apraz refrear as desregradas paixões da mocidade (APULEIO, 1956, p. 79). Psiquê é a mais bela e a última das irmãs, com as quais rompe, significando a ruptura com a família de origem na direção de uma família a ser originada. Começa a história como filha e mortal, termina como mãe e divindade. Pelas mudanças em ambas as personagens, unidas como casal, o mito dialoga diretamente com um tipo de sujeito também orientado pela paixão e pela juventude: os sujeitos leitores em formação. O mito de Eros e Psiquê, assim, corresponde, em mais esse aspecto, a uma história a ser sempre

recontada, mesmo que em um momento no qual a transmissão oral seja menos atuante frente às possibilidades das mídias digitais.

O mito de Eros e Psiquê encontrou várias versões na modernidade, destinadas a todas as faixas etárias<sup>6</sup>. A talvez mais recente, de Guto Lins, *Eros & Psiquê.com.br*, contudo, inova ao incorporar, à releitura da fábula, elementos que aludem às distintas linguagens e mídias por uma reinterpretação ativa que traz consigo o fetiche pelas coisas "triviais, secundárias, alheias" (CANEVACCI, 2013, p. 41) na justaposição de elementos considerados incompatíveis em um mix de códigos de presença simultânea. Guto Lins é designer e professor, atuando no mercado editorial, fonográfico e de entretenimento. É também autor e ilustrador premiado de diversos livros infanto-juvenis.

Em *Eros & Psiquê.com.br*, a narrativa é introduzia a partir de Afrodite, que nasce com um nome e um destino e que foi, na maternidade, "colo fotográfico de todas as enfermeiras, médicos e funcionários do hospital" (LINS, 2011, p. 5). Sua vida é narrada – em forma de fotonovela – da carreira como modelo consagrada ao nascimento do filho Eros. Dotado de grande beleza, seu filho tem, já crescido, o "toque divino de juntar as pessoas" (LINS, 2011, p. 13). Por tal capacidade, mãe e filho abrem um negócio, *Afrodite Recados & Pecados*, uma empresa familiar, com fins lucrativos, "formadora de casais e eventuais famílias". Cabe a Eros, como Diretor de Relacionamentos da empresa, a função de controlar e ampliar o sistema de comunicações para os encontros de amor.

Certa vez, como se registra em uma câmera indiscreta que mantinha ligada em seu quarto, Afrodite consulta um espelho, confidente e conselheiro, quanto à superioridade de sua beleza. Como a rainha má dos contos de fadas e das mídias de massa, em certo momento é alertada sobre a existência de Psiquê. Eros, "filho dedicado e

<sup>&</sup>lt;sup>6</sup> Uma das mais belas releituras do mito é *Psiquê*, de Angela Lago, belissimamente ilustrada e publicada em 2009 pela Cosac Naify.

empresário preocupado" tenta auxiliar a mãe promovendo uma festa para aproximar a rival materna de um outro rapaz. Para evitar ser visto, pelos perigos da própria beleza, o Amor decide promover um baile de máscaras. O artifício, porém, é inútil. Eros se apaixona pela moça. Dançando com ela, impõe a condição de jamais permitir ter seu rosto visto, mas Psiquê aproveita um descuido e desmascara o pretendente. Em um momento em que o texto incorpora, na formatação – na excitação das linhas e das páginas -, a perturbação de Eros, o apaixonado foge do baile, sendo ansiosamente procurado por Psiquê, em uma nova referência às fadas e à indústria cultural, quando o baile, em lugar de unir, separa os enamorados. A moça vai, então, à morada de Afrodite, onde está Eros, e recebe mensagens em QR Code. Todas se relacionam às provas do mito de Psiquê, com destaque a última, que alerta para os perigos em torno da beleza. A página seguinte ao QR Code apresenta uma dobradura em forma de caixa da qual sai uma serpente. Psiquê, na obra de Lins, como no mito, foi traída pela própria desmedida. Por fim, desfalece e é acordada por um beijo cinematográfico do Amor. Na sede do conglomerado empresarial Olimpo, ela é apresentada como novo membro da família e das empresas. Entretanto, para que seu rosto não viesse a apaixonar outro homem futuramente, apenas sua voz, na rádio Psiquê, a ser inaugurada, poderia ser acessível aos meros mortais.

A obra de Guto Lins não apresenta apenas uma atualização do mito. O jogo mixado de códigos, ampliado à justaposição de referências a distintos gêneros e suportes faz de *Eros & Psiquê.com.br* uma manifestação de amplo fetichismo visual no qual se encontram nexos entre distintas referências ligadas a fluxos comunicacionais diversos. O próprio livro é um objeto artístico, que ultrapassa a materialidade utilitária do códex e da paginação impressa. Encadernado com linhas, precisa ser "descosturado" no lado direito para ser aberto à leitura. A capa representa uma máscara, com letras vazadas que fazem o livro "olhar" para o exterior. Aqui o sujeito leitor é como que lido pela história, no que avança nas páginas seguintes quando o papel prateado simula ainda um espelho. Na contracapa, em torno do título da obra, a

palavra "livro" se associa a amor, a sedução, a conquista, a jogo, ao real e ao virtual, tudo acima de um significativo ponto de interrogação. O que, afinal, olha para o leitor enquanto tem o objeto em mãos? O formato físico da obra, inusitado, permite que, previamente, num passar de olhos e de mãos, se apresente a diversidade de recursos e de materiais – folhas de naturezas, texturas e espessuras diversas, letras em diferentes formatos e tamanhos, ilustrações em vários estilos, adesivos, páginas destacáveis, atividades e desafios – entre diferentes gêneros textuais que a compõem, num processo narrativo não linear, transformando a obra em objeto de arte cujo conteúdo a ser lido, o texto, é tão importante e legível quanto seu suporte físico...

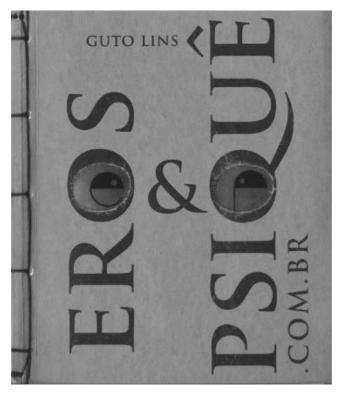


Figura 2. Capa de Eros & Psiquê.com.br Fonte: Eros & Psiquê.com.br



Figura 3. Detalhe da contracapa de *Eros & Psiquê.com.br*Fonte: *Eros & Psiquê.com.br* 

No livro, seja na encadernação, seja no texto, é constante a ideia de espelhamento, alusão à beleza e também, implicitamente, ao reflexo da imagem da qual se desdobra uma alteridade, como uma representação desfigurada de outro, do original. Tal noção aparece na disposição textual que obriga a uma leitura perante um espelho e faz ainda necessário que o livro gire de ponta-cabeça para ser lido. É o caso da página, de numeração refletida, em que se narram as consultas de Afrodite ao espelho.

52 -Espelho, espelho meu, existe alguém no mundo mais bela do que eu? E o espelho que era mágico e não era bobo: -Ontem, hoje e no dia que virá, tu és a mais bela e sempre serás! saylou nun a Esse ritual matinal secreto a deixava ainda mais jim 10d bela e pronta para o dia-a-dia de trabalho. Nada como um espelho para colocar as coisas em cujo jour sens quiqos lugares, lobo mau... se pintar algum tapete mágico, sore some sew aparece bruxa, conced certa manhă incerta, Afrodite, após o banho n inpad ... Dadii matinal com água de coco e folhas de hortelã, Ih... espelho mágico? Espelho, espelho meu... já vi esse

Figura 4. Texto em espelho Fonte: Eros & Psiquê.com.br

As personagens de Lins, em si, são reflexos declaradamente distorcidos. Como elemento metatextual bidimensional, de forma similar a uma janela aberta no decorrer do texto em uma relação espaço-informação, a obra esclarece sobre os deuses "originais", em uma paginação específica. Para explicar quem são os personagens mitológicos Psiquê, Eros e Afrodite "originais", o texto abre o livro na vertical, lembrando a leitura em rolo, sem extrair do texto as marcas do digital, deixando

entre as palavras marcas similares às de parágrafo e símbolos ocultos de formatação de documentos escritos no computador.

EROS.·O·ORIGINAL.·ERA·O·DEUS·GREGO·DO·AMOR.

SEGUNDO·SUA·LENDA·MAIS·FAMOSA.·EROS·ERA·FILHO·DE

AFRODITE·COM·ARES·E·VIVIA·PELO·MUNDO·A·DISPARAR·FLECHAS

DE·AMOR·ENTRE·MORTAIS·E·IMORTAIS.·¶·O·DEUS·TORNAVA

INSENSATOS·OS·CORAÇÕES·QUE·INFLUENCIAVA.·LEVANDO

AFORTUNADOS·APAIXONADOS·A·UMA·BUSCA·INCANSÁVEL

PELO·AMOR·VERDADEIRO.·¶·MAIS·FAMOSO·EM·SUA·VERSÃO

ROMANA·(CUPIDO:·O·MENINO·TÃO·AMADO·QUANTO·TEMIDO)

POSSUÍA·A·GRAÇA·DE·SER·UM·DOS·MAIS··IRRESISTÍVEIS·E·BELOS

DEUSES·DO·OLIMPO.·DESPERTOU·INTENSAS·E·BELAS·PAIXÕES.·SENDO

MAIS·CONHECIDA·SUA·HISTÓRIA·DE·AMOR·COM·A·MORTAL·PSIQUÊ.

Figura 5. Uma "janela" sobre Eros, o "original". Fonte: Eros & Psiquê.com.br

Da mesma forma, dentre outros elementos metatextuais, como receitas, comentários, informações adicionais, a obra, quando narra o destino de Psiquê, afirma, nas entrelinhas, um encaminhamento de leitura em síncrese diásporica:

Psiquê significa a personificação da alma, o sopro da vida e a necessidade vital de transformação. Crenças gregas populares relacionavam à alma a imagem de uma borboleta, o que fez com que nossa deusa fosse frequentemente representada com asas (Lins, 2011).

A alma aqui é diásporica e vagante em um eu em metamorfose. A subjetividade de Psiquê ou da psique contemporânea não se afirmam em uma identidade sedentária e imóvel, compacta e única. Essa unidade "se desmancha por uma afirmação de possíveis pluralidades identitárias no interior do mesmo sujeito" (CANEVACCI, 2013, p. 236). Não por acaso, a *SincrétiKa*, de Canevacci, abre-se com uma epigrafe do

Decameron que trata justamente do amor<sup>7</sup>. A "vaga luz" significa uma beleza que apaixona e desorienta, obrigando a uma viagem sem meta. Essa parece ser a nova dimensão da arte e dos sujeitos: vagar para além dos limites, não se sujeitar às imposições, refazer os caminhos antes interditados. Nesse sentido, a arte em si não se basta em apenas uma produção finita, mas se expõe como reflexo perante outros espelhos deformadores. Não bastando em si, o livro Eros & Psiquê.com.br, suporte, mídia autorreferente, metatextual, se estende à internet, com mais páginas que se desdobram, por força de mais de uma provocação nas páginas impressas, no digital, no site www.erosepsique.com.br. Nesse ambiente, há jogos e curiosidades, as quais animam um pouco mais a releitura do mito clássico, ao mesmo tempo parecendo reforçar o quando nossa identidade é feita de costuras e descosturas. Ali, em um dos links se afirma: "Nossa história é uma colcha de retalhos construída com vários saberes somados". Somos líquidos, multivíduos vagando em pluriversos juvenis:



Figura 6. Curiosidades e retalhos Fonte: erosepsique.com.br

7

<sup>&</sup>lt;sup>7</sup> Amor, la vaga luce, che move dà begli occhi di costei, servo m'ha fatto di te e di lei.

### Considerações finais, por ora...

A Jornada de Literatura de Passo Fundo, desde os momentos iniciais de sua organização, mesmo sob pena de às vezes ser mal compreendida, jamais buscou refugiar-se na especificação de um campo, na qualificação definida do que poderia ser ou não literário. Sua preocupação sempre foi o diálogo entre a literatura e a cultura, ciente do quanto tais termos não têm conceito fixo ou compreensão definitiva. No fluxo transformador dos sujeitos e da arte, a compreensão demasiada tem em si o ônus de uma esterilidade intrínseca. Segundo Canevacci:

A verdade dos pensamentos está na sua capacidade de antecipar o processo de construção social-cultural da realidade, no qual é possível "sentir" o que antes parecia incompreensível e que agora pode ser visivelmente anunciado. O pensamento que "compreende" não antecipa, não vê, não pensa o não-dito. Esse pensamento se conforma diante da realidade: não tem visões. Tarefa do pensamento verdadeiro – na sua mais descentrada manifestação racional – é de ser incompreensivelmente visionário (2013, p. 238).

As relações entre literatura e cultura, indispensáveis para quem pensava, desde sempre em leitura, vieram acrescentar, da parte cada uma desses termos em contato, novos pontos de aproximação e novos termos indissociáveis. Da defesa dos livros de papel nos anos 80 à vigência da portabilidade e da virtualidade de uma biblioteca digital em um telefone celular, o percurso de toda uma história buscava e busca pensar o "não-dito". Quando as primeiras discussões, praticamente ao fim da Jornada de 2011, decidiram pelo jovem como tema da edição de 2013, alguma incompreensão visionária acenava para o que aconteceria meses depois, em um junho incendiário nas ruas das grandes metrópoles e para o que aconteceria após, em 2014, com as grandes (e indesejadas) concentrações nos shoppings do país. "Algo nos dizia", poderíamos alegar... Mas o certo é que sabíamos que pensar a

tecnologia sem pensar no sujeito que opera e a recria é imaginar que poderia haver uma literatura sem leitores (o que já se pensou, por sinal, de alguma forma, nos círculos mais restritos dos estudos literários). Sujeitos e artes são elementos especulares que se ressignificam trocando "olhares recíprocos", quanto mais em uns tempos híbridos. Alckmar dos Santos, refletindo sobre a identidade no ciberespaço, elabora uma das imagens mais provocativas das relações entre sujeito e tecnologia:

Por paradoxal que pareça, uma experiência importante que podemos ter dos textos eletrônicos ocorre justamente quando desligamos o computador e se apaga a tela. Nesse fundo opaco, que instantes atrás eram brilhos e pixels, aparece uma figura esvanecente, nossa fisionomia, um pálido reflexo que somente se mostra a partir do monitor desligado. Desligada a máquina, o que se vê ao fundo, precariamente refletida, é então essa nossa imagem diante da tela, trazendo à tona e explicitando, talvez, o incômodo de uma posição em que nos surpreendemos inquirindo subjetividades e perturbando identidades. É como se se reproduzisse a difícil posição do indivíduo que na Procura da Poesia, de Carlos Drummond de Andrade, se vê colocado diante da palavra, que "te pergunta, sem interesse pela resposta, / pobre ou terrível que lhe deres: / Trouxeste a chave?". Contudo, o que perturba e incomoda é que o inquisidor não é palavra alguma, ele se parece muito conosco! (SANTOS, 2003, p. 24).

A tela do computador nos devolve esse outro que passa a ser parte perturbadora de nós, da mesma forma como o olhar *de Eros & Psiquê.com.br* nos pergunta, atrás de si, na contracapa, o que é – um livro, um jogo? Feito de coisas diferentes que não se negam distintas, costurada de diferenças que nunca se sintetizam, a obra não se diz literatura nem livro, nem arte nem entretenimento/passatempo, mas tudo isso e mais, ao mesmo tempo e em lados e laços distintos – para falar mais sobre "nós" (pronome) e "nós" (substantivo).

O jovem do presente líquido tornou-se o representante desse novo tempo, ou melhor, dessa nova temporalidade. Nela, fragmentos glocais indicam-se não como restos de algo ferido, arruinado, mas como frações livres, vagas, pelas quais é razoável afirmar que o provisório terá a imortalidade e a permanência dos deuses, e que a juventude, sempre nova e presente, nos alertará, constantemente, de que tudo que está pronto... ainda está por se fazer.

#### REFERÊNCIAS

APULEIO, Lucio. *Amor e psique*. Tradução: Paulo Rónai e Aurélio Buarque de Holanda. Rio de Janeiro: Civilização Brasileira, 1956. 87 p.

CANEVACCI, Massimo. *Culturas eXtremas*. Mutações juvenis nos corpos das metrópoles. Rio de Janeiro: DP&A, 2005. 200 p.

\_\_\_\_\_. *SincrétiKa*. Explorações antropológicas sobre artes contemporâneas. São Paulo: Studio Nobel, 2013.293 p.

FEENBERG, Andrew. *Transforming technology*: a critical theory revisited. Oxford: Oxford Press, 2002. 220 p.

LÉVY, Pierre. Cibercultura. São Paulo: Ed. 34, 1999. 264 p.

\_\_\_\_\_. *A inteligência coletiva*: por uma antropologia do ciberespaço. 5. ed. São Paulo: Loyola, 2007. 214 p.

LINS, Guto. *Eros & Psiquê.com.br*. Rio de Janeiro: Manifesto Design, 2011.

SANTOS, Alckmar. *Leitura de nós*: ciberespaço e literatura. São Paulo: Itaú Cultural, 2003. 148 p.

# A narrativa transmídia como estratégia de incentivo à leitura

Elisa Cristina Delfini Corrêa<sup>1</sup>

Narrar histórias é uma prática que acompanha o homem desde o início da sua vida em sociedade. Desde as comunidades mais primitivas, as histórias contadas de pais para filhos tornaram vivas e ajudaram a preservar a cultura e a identidade dos povos. As imagens nas cavernas, que correspondem aos primeiros registros da comunicação escrita entre os seres humanos, já deixaram evidentes relatos da atividade do homem primitivo, contando e preservando suas histórias de caça e descobertas até os dias de hoje. Nessas sociedades, o contador de histórias possuía um importante papel social, pois detinha as experiências e a sabedoria de sua época, trabalhando com a construção oral coletiva que se fundamentava na identidade cultural do seu povo.

Séculos depois, contar histórias permanece ainda uma atividade presente em todo o mundo e é considerada uma ferramenta bemsucedida de incentivo à leitura. São muitas as pesquisas que apontam para os benefícios da contação de histórias para crianças, adolescentes, jovens e adultos, cooperando para o bem-estar físico e emocional dos ouvintes. Dentre os principais benefícios, fica evidenciada principalmente a sua capacidade de instigar as pessoas a adentrar o universo dos livros e assim ampliar sua visão de mundo, contribuindo para o crescimento individual e inclusão social.

<sup>&</sup>lt;sup>1</sup> Professora do Programa de Pós-Graduação em Gestão da Informação da Universidade do Estado de Santa Catarina (UDESC). CONTATO: correa.net@iq.com.br

A simplicidade do momento em que alguém se dispõe a ler um trecho de uma pequena história já traz em si um encantamento capaz de despertar no ouvinte a sua criatividade, levando-o a sentir o desejo de experimentar mais e conquistar sua autonomia nessa busca. Ramos (2011, p. 27) destaca o valor dessa interação, afirmando que

no caso da experiência da contação de história, as palavras proferidas pelo contador são como linhas desenhadas pelo ar. Enquanto o contador liberta as palavras presas no texto, o ouvinte, leitor indireto do texto narrado, vai criando e interpretando os desenhos, adentrando-se em um mundo mágico e tornando-se co-autor da história.

Por outro lado, contar bem uma história exige do contador todo um "pacote" de técnicas e habilidades, construindo um cenário que o conduzirá à *performance* ideal para atingir o sucesso desejado. Nas comunidades mais antigas e, talvez ainda hoje nas mais afastadas da chamada 'civilização', as histórias eram contadas em um cenário especial que representava uma espécie de ritual, onde todos se assentavam ao redor de uma fogueira à noite, por exemplo, criando o clima ideal.

Além disso, há a necessidade de uma postura corporal e vocal por parte do contador que sejam capazes de chamar e prender a atenção dos ouvintes durante a história. Para isso, pode-se recorrer também ao uso de objetos, figuras e sons. Contar histórias, portanto, "é arte performática, na qual se busca compartilhar vivências através da voz, do corpo e dos gestos" (SCHERMACK, 2012, p. 4).

Artefatos diversos também são úteis nessa *performance*: um chapéu, um fantoche ou qualquer outro objeto "misterioso" são poderosos instrumentos na hora de ilustrar o texto narrado.

Assim, observa-se que para narrar uma história utiliza-se um conjunto de artefatos que dão vida ao texto e levam os ouvintes a interagir e

mergulhar no universo apresentado pelo contador. Todos esses aparatos são, na verdade, recursos utilizados a fim de potencializar a história, aumentando ainda mais o poder de imaginação e criatividade tanto de quem a conta, quanto de quem a ouve.

Os contadores de história do século XXI possuem ao alcance das mãos recursos nunca antes imaginados em séculos anteriores, aliando as tecnologias e mídias atuais à arte de envolver e interagir com seu público. Schermack (2012, p. 8) afirma que:

A contação de histórias, no início do século XXI, envolve um encontro entre a narração oral tradicional e o suporte digital na medida em que a tecnologia informatizada (a luz da tela do computador) traz a necessidade do uso de novos suportes para concretizar uma das artes mais antigas: a contação de histórias. A facilidade com que os jovens manipulam o computador sugere que se considere a linguagem do meio digital (como por exemplo, os hipertextos, as imagens coloridas, músicas, vídeos, etc) como um instrumento importante no contexto educacional, principalmente para a formação de leitores críticos.

Levando em consideração a facilidade de uso de ferramentas tecnológicas por crianças, adolescentes e jovens, bem como toda consequente capacidade de interação que este público possui nesse ambiente, é necessário unir o tradicional e o digital à atividade de narrar histórias a fim de torná-las um instrumento ainda mais poderoso na formação de leitores no contexto atual.

Trata-se, portanto, de proporcionar uma convergência entre as mídias conhecidas, oportunizando um novo caminho ainda mais instigante e participativo ao mundo do encantamento e do conhecimento que advém de uma história bem contada.

Nesse contexto de convergência, surge a narrativa transmídia como modalidade de contação de histórias que, aproveitando diferentes

canais nos quais estas se desenvolvem, oportunizam e potencializam a participação de todos em um ambiente interativo.

Para entender melhor esse conceito e refletir sobre sua potencial contribuição para o desenvolvimento do gosto pela leitura, este texto pretende introduzir a definição de narrativa transmídia, discutindo como essa estratégia pode ser utilizada a fim de impulsionar a formação de leitores, apresentando para isso, alguns exemplos práticos de sua utilização.

# Narrativa transmídia: o que é

Segundo Gosciola e Versuti (2012), o termo ainda está em construção, embora em fase conclusiva:

Como conceito, a narrativa transmídia vem se definindo aos poucos e está bem próxima de sua conclusão. Mas os primeiros usos do termo são muito significativos (Gosciola, primeiramente 2011). Aparece como trans-media composition (Welsh, 1995, p. 97) definido pelo compositor e instrumentista Stuart Saunders Smith, para a peça Return and Recall de 1976, como uma composição em que melodia, harmonia e ritmo diferentes de cada instrumento são novamente compostos por cada executor, que complementaria a obra em coerente harmonia e sincronia com os outros instrumentistas/compositores da peça (Sauer, 2009).

O conceito transmídia foi introduzido como objeto de estudos no campo da comunicação através da publicação de um livro² cujo termo "transmedia intertextuality" apontava a criação de um sistema de entretenimento no qual crianças experimentavam "a ampliação narrativa de uma história ao assistir à série de desenhos na TV, brincar

-

<sup>&</sup>lt;sup>2</sup> Playing with Power in Movies, Television, and Video Games: From Muppet Babies to Teenage Mutant Ninja Turtles, de autoria de Marsha Kinder

com os brinquedos e ao assistir a um filme no cinema das mesmas personagens" (GOSCIOLA; VERSUTI, 2012).

Contudo, o termo **narrativa transmídia** recebe sua primeira definição no artigo *Transmedia storytelling*, de autoria de Henry Jenkins (2003). No texto, Jenkins afirma que o futuro do entretenimento comercial estaria na capacidade de movimentar as histórias e seus personagens em uma narrativa que pudesse convergir livros, jogos e filmes, fazendo fluir o seu conteúdo através de múltiplas plataformas.

O autor analisa que essa convergência se faz necessária em um contexto social no qual o computador é parte integrante do dia a dia de boa parte da população mundial e cujo público leitor assume características diferenciadas, afirmando que:

Os consumidores mais jovens tornaram-se caçadores e coletores de informação, possuem prazer em rastrear as origens das personagens e pontos da trama e fazer conexões entre textos diferentes dentro da mesma franquia. E, além disso, todas as evidências sugerem que os computadores não anulam outros meios de comunicação, em vez disso, os proprietários de computador consomem, em média, significativamente mais televisão, filmes, CDs, e mídia relacionadas do que a população em geral (op.cit., p.1).

Desse contexto surge o conceito de narrativa apresentado por Jenkins, que une as diferentes mídias disponíveis a fim de contar uma grande história através de um novo modelo de co-criação que utiliza o potencial único de cada uma das diferentes mídias:

Na forma ideal de narrativa transmídia, cada meio faz o que faz de melhor para que a história possa ser introduzida num filme, expandida através de televisão, romances e quadrinhos, e seu mundo pode ser explorado e experimentado também através de jogo. Cada entrada da franquia precisa ser bastante auto-suficiente para permitir o

consumo autônomo. Ou seja, você não precisa ter visto o filme para desfrutar do jogo e vice-versa (p. 2).

O texto é voltado a uma perspectiva comercial da indústria de entretenimento, porém a ideia de contar uma história através de múltiplas mídias possui a especial vantagem de ampliar o leque de público, princípio que também pode ser aplicado à contação de histórias de maneira geral:

A leitura através da mídia sustenta uma profundidade de experiência que motiva mais consumo.[...] Tal abordagem multifacetada para contar histórias vai permitir um modo mais sofisticado, complexo, e gratificante de fazer surgir a narrativa para além das limitações de entretenimento comercial. Mas isso também faz sentido economicamente falando. Diferentes meios para atrair diferentes nichos de mercado. Filmes e programas de televisão, provavelmente possua audiência mais diversificada do que quadrinhos e jogos. Uma boa franquia transmídia atrai um público mais amplo, lançando o conteúdo de forma diferente em diferentes mídias. Se cada trabalho individualmente oferece novas experiências, então um mercado transmídia irá expandir o potencial bruto dentro de qualquer mídia individual (p. 3).

Em sua obra *Cultura da convergência* (2009), Jenkins explora a franquia *Matrix*<sup>3</sup> como exemplo de narrativa transmídia que, segundo o autor, "refere-se a uma nova estética que surgiu em resposta à convergência das mídias – uma estética que faz novas exigências aos consumidores e depende da participação ativa de comunidades de conhecimento. A narrativa transmídia é a arte da criação de um universo" (2009, p. 49).

<sup>&</sup>lt;sup>3</sup> A franquia *Matrix* refere-se a uma trilogia de filmes cinematográficos cuja história, de ficção científica, na qual o mundo real na verdade é apenas uma ilusão criada e comandada por uma central artificial e inteligente que controla os seres humanos. A franquia consta de 3 filmes (*Matrix*, *Matrix Reloaded* e *Matrix Revolution*), 9 desenhos animados (*Animatrix*) e um game (*Enter the Matrix*). Fonte: Wikipedia

Nesse universo, a participação dos "consumidores" da narrativa (que possuem as características já descritas acima) participam

perseguindo pedaços da história pelos diferentes canais, comparando suas observações com as de outros fãs, em grupos de discussão online, e colaborando para assegurar que todos os que investiram tempo e energia tenham uma experiência de entretenimento mais rica (idem).

Sobre a franquia *Matrix*, Jenkins afirma que nunca antes o consumidor de filmes havia sido tão exigido. Já o pré-anúncio do primeiro filme lançou a provocativa pergunta "O que é *Matrix*", instigando o público a buscar respostas na Internet, levando-o a utilizar uma mídia diferente logo no primeiro contato. Durante as sequências, os cineastas "plantaram pistas que só farão sentido quando jogarmos o game. Abordam uma história paralela, revelada por uma série de curtas de animação que precisam ser baixados da web e vistos num DVD separado. Os fãs saíram correndo dos cinemas, pasmos e confusos, e se plugaram nas listas de discussão na Internet, onde cada detalhe era dissecado e cada interpretação possível, debatida (JENKINS, 2009, p. 137).

O conceito de narrativa transmídia é então redefinido no livro de 2009, onde o autor acrescenta a possibilidade do "leitor" experimentar a história também através de uma atração num parque de diversões e explica a autonomia que existe em cada um dos acessos à franquia: "cada acesso à franquia deve ser autônomo, para que não seja necessário ver o filme para gostar do game e vice-e-versa. Cada produto determinado é um ponto de acesso à franquia como um todo." (JENKINS, 2009, p. 138).

Desta forma, o autor afirma que *Matrix* é um entretenimento criado para a era da convergência, pois trata-se de uma história ampla demais para ser contida em uma única mídia e, por isso, utiliza-se de diferentes

textos espalhados em diferentes canais. Trata-se também de entretenimento para a era da inteligência coletiva, citando Pierre Lèvy em sua tese de que a diferença entre autores e leitores, produtores e espectadores, criadores e intérpretes seria dissolvida nesse novo contexto social.

Jenkins concorda que poucas franquias (ou talvez nenhuma ainda) já atingiram o potencial estético da narrativa transmídia. O fato de Matrix utilizar essa estratégia não significa necessariamente que tenha sido perfeitamente aproveitada em cada canal e como um todo. Por outro lado, também conclui que ainda não existem bons critérios estéticos para avaliar essas histórias e que, por isso, ainda trata-se de um campo de muitas incertezas, especialmente porque ainda não são muitas as experiências conhecidas.

Por este motivo, também faz-se necessário deixar claro que nem todos os tipos de narrativa que utilizam mídias diferentes podem ser consideradas transmidiáticas. Por ser um conceito bastante recente e ainda em construção, é comum encontrar alguns supostos usos da estratégia que, na verdade, não correspondem exatamente às ideias acima descritas.

Portanto, narrativa transmídia não é, em primeiro lugar, narrativa **multimídia**. Contar uma história utilizando textos, imagens e sons em um mesmo canal, não representa absolutamente uma narrativa transmídia. Nesta, existe uma dispersão de conteúdos em diferentes plataformas, completando uma grande história.

Em segundo lugar, também não é narrativa **crossmídia**. Esta representa geralmente um projeto publicitário que divulga uma mesma história em diferentes plataformas. O termo vem do inglês cross que significa "cruzar" ou "atravessar" a mídia, referindo-se ao fato de utilizar mais de uma mídia para levar o conteúdo, que é a mensagem que:

é distribuído através de diferentes mídias (o meio) para atingir o público (o receptor), mas tudo isso acontece sem que a mensagem tenha qualquer alteração de um meio para outro. O sentido básico desse termo é que uma pessoa possa acessar o mesmo conteúdo por diferentes meios" (ALTERMANN, 2013).

Para Altermann, o termo *transmedia* (que também vem do inglês) significa "além da mídia, ou seja, o conteúdo sobressai a mídia" (idem). O autor faz uma analogia com a narração de um jogo de futebol, deixando a diferença bem clara:

#### TRANSMÍDIA

- O jogo é transmitido na TV, no rádio e na internet;
- Na TV o jogo é transmitido ao vivo com narração do comentarista;
- No rádio são transmitidos os bastidores do jogo, como entrevista com torcedores, exjogadores, etc.;
- Na internet a pessoa pode conferir todas as substituições, cartões, gols, ver replays, históricos de partidas de ambos os times, etc;
- Resumo: Os meios, juntos, fortalecem uns aos outros e complementam a sua mensagem.

#### **CROSSMEDIA**

- O jogo é transmitido na TV, no rádio e na internet;
- Na TV o jogo é transmitido ao vivo com narração do comentarista;
- No rádio o jogo é transmitido ao vivo com narração do comentarista;
- Na internet o jogo é transmitido ao vivo com narração do comentarista;
- Resumo: Os meios, juntos, levam o conteúdo ao máximo de pessoas

Fonte: Blog Midiatismo, 2013

Algumas profissões como o jornalismo, por exemplo, utilizam princípios da narrativa transmídia para a transmissão de notícias a partir de diferentes canais. O conceito é aplicável à prática do jornalismo que, em sua concepção mais básica, refere-se ao ato de contar histórias. Por

isso, a narrativa transmídia tem chamado a atenção dos jornalistas que vem incorporando aspectos dessa técnica em sua atuação profissional, especialmente chamando à participação o público em geral, que envia vídeos ou fotos que complementam o conteúdo a partir de seus celulares ou equipamentos similares.

No entanto, por ser muito recente e ainda carente de melhor entendimento, é muito provável que nem sempre seja aplicado da maneira mais próxima à da proposta apresentada por Jenkins. A principal razão para isso é que é discutível o fato de que o jornalismo, de maneira geral, possua o real entendimento da narrativa transmídia enquanto processo estratégico de construir e difundir uma determinada história a partir de diferentes mídias e de forma colaborativa, o que fica praticamente inviável no caso dos "furos" jornalísticos, por exemplo.

O traço mais marcante da narrativa transmídia reside na utilização de diferentes mídias na construção e distribuição de uma história (ou um fato jornalístico), onde os conteúdos por sua vez convergem todos à mídia principal, nas palavras de Jenkins, a "nave mãe" (NAVARRO, 2010, p. 12). No entanto, é preciso que sejam feitas duas considerações: em primeiro lugar, o uso de mídias nessa técnica (ou estratégia) não se refere apenas às mídias digitais. O simples fato de usar diferentes plataformas como *Twitter*, *Facebook* ou outras "redes sociais" da Internet não exprime o sentido mais amplo do conceito. Em sua essência, a narrativa transmídia faz convergir igualmente a mídia televisiva, radiofônica, cinematográfica e impressa, além das mídias digitais.

O segundo aspecto a ser considerado refere-se à dinâmica pela qual se constrói uma narrativa transmídia a partir de seis elementos distintos: a história, a experiência, a audiência, as plataformas, o modelo de negócio e a execução (Pratter, 2011<sup>4</sup>). Isto equivale a dizer que é necessário haver um planejamento que parta desde a escolha da história que deverá ser contada, das mídias que serão utilizadas e dos níveis desejados de participação do público. Portanto, assim como o simples uso de diferentes mídias sociais não caracteriza a narrativa transmídia na totalidade de sua concepção, o mesmo pode ser dito sobre o simples fato de haver uma participação espontânea ou solicitada do público em colaborar com a história, com o envio de imagens, por exemplo.

Para Alzamora e Tárcia (2012, p. 31), um jornalismo transmidiático pressupõe "não apenas complementaridade midiática, embora esta seja uma característica relevante do processo, mas principalmente deslocamento das características tradicionalmente marcadas pelos ambientes midiáticos". Como exemplos, são citados o jornalismo de imersão e os *news games*<sup>5</sup>.

Desta forma, discute-se se o conceito de narrativa transmídia pode ou não ser considerado o mais adequado para o tipo de jornalismo 2.0 praticado comumente. Alzamora e Tárcia levantam essa questão e citam outros autores que apontam essa reflexão:

Pernisa Júnior (2010) considera difícil visualizar a prática do Jornalismo Transmidiático. Para este autor, o que se vê, na maioria das vezes, são modelos de veículos analógicos sendo transportados para o meio digital. Ele defende o pensamento de Jenkins como base conceitual, porém ressalta a necessidade de não se prender ao conceito, uma vez que toda

<sup>5</sup> A ideia fundamental do Jornalismo Imersivo seria permitir ao participante entrar em um cenário representativo da história, criado virtualmente<sub>7</sub>. A perspectiva da participação ativa do usuário na construção da narrativa e a possibilidade de compartilhá-la em mídias sociais são, nesse caso, aspectos relevantes em sua caracterização como experimento transmídia.(ALZAMORA; TÁRCIA, 2012, p. 31)

\_

<sup>&</sup>lt;sup>4</sup> PRATTER, Robert. **Getting started in transmedia storytelling**. 2011. E-book: Kindle Store, 2001. Disponível em: < http://www.tstoryteller.com/getting-started-in-transmedia-storytelling>. Acesso em: 25 abr. 2011.

a sua teoria está baseada no mundo do entretenimento e do marketing. O autor enxerga o Jornalismo Transmidiático como uma proposta conceitual, que deve ser vista a partir do ponto de vista da experimentação e não de uma realidade já consumada.

A partir de toda conceituação necessária feita até aqui, somos conduzidos agora à reflexão sobre as possibilidades de uso dessa narrativa como estratégia de incentivo à leitura, o que será feito a seguir.

## Narrativa transmídia e o gosto pela leitura

O que fica claro sobre a narrativa transmídia é que ela tem como base uma grande história, contada através de diferentes mídias que são aproveitadas ao máximo em sua potencialidade, e cujo valor é agregado também a partir da participação do público interessado.

Sendo assim, a primeira e mais evidente contribuição para o incentivo à leitura é conduzir o "ouvinte" à condição de caçador de partes da história a diferentes canais. Assim, o interessado em buscar a informação é levado a ler os livros, assistir os filmes e a jogar os games (como no caso de *Matrix*). Isso por si só já serve como incentivo à diferentes caminhos para a leitura, diversificando os canais e instigando a curiosidade e a criatividade.

Desde atividades simples que podem ser realizadas dentro de uma sala de aula comum, unindo filmes e livros, por exemplo, com criatividade e dedicação por parte do educador, é possível criar uma espécie de "caça à história", envolvendo as estratégias da narrativa transmídia, levando alunos a diferentes espaços da escola, como bibliotecas, quadras de esportes e outros locais a fim de construir a história com a participação dos alunos.

No entanto, contar uma história com essas características exige planejamento. É preciso pensar as mídias a serem utilizadas, distribuir seu conteúdo e conduzir o público a buscar a informação e o entendimento da história em cada canal a fim de construir um diálogo transmidiático capaz de retornar à história original abrigada na "navemãe" (mídia principal).

Este recurso é viável e pode ser aproveitado, por exemplo, na Educação, na Psicologia, na Biblioteconomia ou em qualquer área que considere a contação de história como atividade profissional.

Na área da Educação, Falcão (2011) afirma que a narrativa transmídia deve ser utilizada como instrumento de mediação pedagógica, sendo:

um elemento catalisador no relacionamento entre professor e aluno, servindo como instrumento de mediação pedagógica. Isso se deve ao fato de as narrativas transmídia se basearem em colaborativismo, participação, criação coletiva e expressão em diferentes formatos, linguagens e canais guiados por uma narrativa central".

Uma das formas de fazer essa conexão é através dos processos de fanfictions, que são narrativas criadas por fãs a partir de um universo fictício já existente como Harry Potter, por exemplo. Uma experiência realizada por um grupo português que trabalha com os conceitos de REA (Recurso Educacional Aberto), relatada por Gosciola e Versuti (2012), sugeriu a seguinte estratégia:

com dispositivos móveis conectados à internet sem fio e monitorados por GPS, poderá ser o ponto de partida para a reconstrução imaginária de uma *Hogwarts School of Witchcraft and Wizardry*, na própria escola a ser aplicado o projeto. Os conhecimentos dos alunos, tanto da sua escola quanto da história de Harry Potter, se mesclariam. É uma oportunidade de integração das disciplinas aos repertórios externos à escola e aos recursos tecnológicos... Como

atividade de coaprendizagem podemos propor aos alunos a redação de uma *fanfic* e sua extensão para um projeto transmídia sobre Harry Potter, de forma colaborativa, usando o *google docs*, por exemplo. E podemos solicitar que produzam um vídeo a partir de imagens dos personagens, editando-o de acordo com as novas versões da "estória" criada. Ainda podemos solicitar que recriem imagens licenciadas com *Creative Commons* de acordo com as *fanfics* criadas <sup>6</sup>

Levar o aluno a participar da construção da história é um caminho eficiente para instigar o desejo de saber mais, de conhecer outras histórias, de penetrar no universo ficcional não apenas de forma passiva, mas assumindo características das personagens. Em tempos de web 2.0, todos somos chamados a assumir a co-autoria do material informativo que circula na Internet. Através das *fan-fictions* isso não apenas é possível a crianças e adolescentes, como também oportuniza encontros criativos *off-line*, em feiras como a *Comic Con*, que acontece também no Brasil e reúne fãs de revistas em quadrinhos e séries de TV<sup>7</sup>.

Outra forma de utilizar a narrativa transmídia para incentivar a leitura é através da gamificação, que apesar de tratar-se de uma estratégia amplamente utilizada em empresas para motivação de seus funcionários, também pode (e deve) ser aplicada à leitura. A gamificação de livros consiste na utilização de aplicativos para leitura de livros digitais, que auxiliam na compreensão da história e das ideias contidas, instigando os leitores através de desafios em jogos rápidos e divertidos ao final de cada capítulo. O uso desses aplicativos proporciona também a oportunidade da criação de comunidades de leitores que trocam suas experiências e compartilham anotações de suas leituras. Ou seja, uma espécie de "clube do livro digital".

<sup>6</sup> http://oer.kmi.open.ac.uk/?page\_id=428

http://blogs.estadao.com.br/link/o-que-esperar-da-comic-con-brasileira/

<sup>&</sup>lt;sup>8</sup>Como exemplo, cita-se o app Readsocial (Dopa Labs) http://www.guickdrops.com.br/2013/06/gamificacao-da-leitura.html

No Brasil, já existem experiências semelhantes disponíveis. *A saga Birka*, criada pela Editora Buqui (RS) conta, em cinco volumes, a história ficcional de um mundo controlado por regras extremamente rígidas de comportamento e do retorno de uma personagem que fora banida por questionar tais regras. As vantagens e o objetivo desse tipo de publicação, segundo seus criadores<sup>9</sup>:

Para resgatar leitores que hoje estão em outras mídias e também para enriquecer a experiência do usuário a Editora Buqui inicia projeto inédito, no Rio Grande do Sul, de transmídia de livro para game [...] Também são selecionados os trechos da obra que podem agregar mais experiência ao jogador. O game vira um complemento do livro e as duas mídias se comunicam, para estimular a leitura.

Ao contrário do que se possa pensar, apesar de atingir diretamente o público infanto-juvenil e jovem, as estratégias de gamificação também atingem uma parcela considerável de adultos<sup>10</sup>:

A mecânica de jogo não tem apelo somente para este grupo principal [Geração Y]. Muitos adultos têm se mostrado tão afiados quanto os jovens ao competir e divulgar seus resultados - princípios essenciais da gamificação. Nos Estados Unidos, 37% dos aficcionados por jogos eletrônicos têm mais de 35 anos. No Reino Unido e nos Estados Unidos, adultos que apreciam jogos com fortes elementos sociais são maioria dos jogadores em dispositivos móveis. Isso acontece devido ao predomínio crescente da mídia digital: nos Estados Unidos, a publicidade online já alcança a TV rapidamente e a previsão era de que superaria os anúncios impressos em 2012.

Essa estratégia já é utilizada em algumas bibliotecas, como a Biblioteca Pública de Nova York, que realizou uma noite inteira de jogos com

<sup>10</sup>http://www.accenture.com/br-pt/outlook/Pages/outlook-journal-2013-whygamification-is-serious-business.aspx

<sup>9</sup> http://www.planetaempreendedor.com.br/?p=9228

adolescentes e jovens num evento chamado *Find the Future* (Encontre o futuro)<sup>11</sup>:

O game foi criado para incentivar os jovens a descobrirem seu próprio futuro, colocando-os cara a cara com objetos e livros relacionados a pessoas que fizeram uma diferença extraordinária no mundo", explica Jane no site do evento. Baseados em 100 objetos que "inspiraram a humanidade", os participantes selecionados vão ajudar a escrever "100 histórias que nunca foram contadas.

Segundo pesquisa realizada pela NPR<sup>12</sup>, esse tipo de atividades em bibliotecas aumenta a permanência das pessoas nesse ambiente e a circulação de livros em torno de 15 a 20%<sup>13</sup>.

## Considerações finais

Os benefícios sociais da formação de leitores são indiscutíveis e no Brasil, especialmente, ainda existe uma imensa lacuna a ser preenchida nesse campo. A narrativa transmídia apresenta-se como uma alternativa viável de incentivo à leitura, unindo diferentes canais bastante difundidos e utilizados por crianças, jovens e adultos.

Um campo ainda inexplorado e novo que, se levado com a seriedade que merece, pode trazer resultados bastante compensadores para uma transformação cultural necessária no país.

As ferramentas estão ao alcance das nossas mãos. Convido você a experimentá-las e a tornar-se não apenas um leitor melhor e mais

http://www1.folha.uol.com.br/folhateen/899787-jovens-participam-de-game-noturno-na-biblioteca-de-nova-york.shtml

<sup>&</sup>lt;sup>12</sup> NPR (National Public Radio), Organização de mídia sem fins lucrativos da América.

<sup>&</sup>lt;sup>13</sup>http://www.baixakijogos.com.br/noticias/games-em-bibliotecas-podem-incentivar-a-leitura-diz-pesquisa\_96925.htm

conectado mas, na medida do possível, a tornar-se um agente de mudança social no contexto onde você está inserido.

## **REFERÊNCIAS**

ALTERMANN, D. Vamos entender a diferença entre crossmedia e transmídia. Disponível em: http://www.midiatismo.com.br/o-marketing-digital/vamos-entender-a-diferenca-entre-crossmedia-e-transmidia-cirandablogs. Acesso em: 07 fev.2014

ALZAMORA, G.; TÁRCIA, L. Convergência e transmídia: galáxias semânticas e narrativas emergentes em jornalismo. *Brazilian Journalism Research*, Brasília, v.8, n.1, 2012. Disponível em: http://bjr.sbpjor.org.br/bjr/article/view/401/370 FALCÃO, B. *A narrativa transmídia como instrumento de mediação pedagógica*: o papel do professor no contexto das novas tecnologias. Blog Professora Bárbara Falcão, mai./2011. Disponível em: http://professorabarbarafalcao. wordpress.com/2011/05/31/a-narrativa-transmidia-como-instrumento-demediacao-pedagogica-o-papel-do-professor-no-contexto-das-novas-tecnologias/

GOSCIOLA, V.; VERSUTI, A. (2012). Narrativa transmídia e sua potencialidade na educação aberta. In: Okada, A. (Ed.) (2012) *Open Educational Resources and Social Networks:* Co-Learning and Professional Development. London: Scholio Educational Research & Publishing. Disponível em: http://oer.kmi.open.ac.uk/?page\_id=428. Acesso em: 06 fev.2014.

JENKINS, H. Transmedia storytelling. *MIT Technology Review*. Disponível em: http://www.technologyreview.com/news/401760/trans media-storytelling/. Acesso em: 05 fev.2014.

\_\_\_\_\_. Cultura da convergência. 2. ed. São Paulo: Aleph, 2009.

NAVARRO, V. Sites of convergence: an interview with Henry Jenkins. Niterói, *Contracampo*, n.21. ago./2010 Disponível em: http://www.uff.br/contracampo/index.php/revista/article/view/77/57. Acesso em: 05 fev.2014

PRATTER, R. *Getting started in transmedia storytelling*. 2011. E-book: Kindle Store, 2001. Disponível em: < http://www.tstoryteller.com/ getting-started-in-transmedia-storytelling>. Acesso em: 08 fev. 2014.

RAMOS, Ana Claudia. *Contação de histórias*: um caminho para a formação de leitores? Dissertação Mestrado em Educação, Universidade Estadual de Londrina, 2011. Disponível em: http://www.uel.br/pos/mestredu/imaqes/

stories/downloads/dissertacoes/2011/2011\_-\_RAMOS\_Ana\_Claudia.pdf. Acesso em: 05 fev.2014.

SCHERMACK, Keila de Quadros. A contação de histórias como arte performática na era digital: convivência em mundos de encantamento. In: CONGRESSO INTERNACIONAL DE LEITURA E LITERATURA INFANTIL E JUVENIL, 3., FORUM LATINO AMERICANO DE PESQUISADORES DE LEITURA, 2. Porto Alegre, 2012. *Anais*. Porto Alegre, Edipucrs, 2012. Disponível em: http://ebooks.pucrs.br/edipucrs/anais/IIICILLIJ/Trabalhos/Trabalhos/S10/keilaschermack.pdf. Acesso em: 04 fev.2014.

# Sobre a paratextualidade: um relato de teoria aplicada em sala de aula

Cláudio Augusto Carvalho Moura<sup>1</sup>

# Introdução

O presente texto constitui-se de um relato ensaístico sobre um episódio bem sucedido de Teoria Literária aplicada em sala de aula de graduação ocorrido há alguns anos. Trata-se, na verdade, de uma de minhas primeiras experiências docentes no ensino superior, que se deu através do cumprimento do estágio de docência exigido aos bolsistas da Coordenação de Aperfeiçoamento de Pessoal do Ensino Superior (CAPES). No meu caso, à época, discente do Mestrado Acadêmico em Letras/ Estudos Literários da Universidade Federal do Piauí (UFPI).

Como o leitor já deve ter observado, optei pela utilização da primeira pessoa em nome de uma fluência melhor em minha escrita. Assim o fiz também por achar mais sincero de minha parte, uma vez que este texto não está sendo escrito com o intuito de apenas teorizar acerca da paratextualidade, mas sim de narrar essa experiência e suas implicações. Quanto à teoria, a mesma é aqui utilizada como parte da narrativa tal apresentada em sala à época, sem acrescentar nada referente às leituras posteriores que fiz, na intenção (mesmo que inocente) de tentar me manter, o máximo possível, fiel ao ocorrido.

\_

<sup>&</sup>lt;sup>1</sup> Professor do Departamento de Letras da Universidade Federal do Piauí (UFPI), doutorando do Programa de Pós-Graduação em Literatura da Universidade Federal de Santa Catarina (PPGL/UFSC), membro do Núcleo de Pesquisa em Literatura Digitalizada (NUPLID/UFPI) e do Núcleo de Pesquisa em Informática Literatura e Linguística (NuPILL/UFSC). CONTATO: ca.carvalho@ufpi.edu.br

Outro ponto sobre este texto seria o porquê da demora em trazer-lhe à baila. A explicação para tanto é simples: o material necessário para escrever este relato julgava-se perdido. A experiência, contudo, permaneceu sempre viva em minha cabeça, me martirizando o fato de saber que eu não escreveria sobre se não tivesse como mostrar exatamente aquilo do que eu falava. Simplesmente não poderia fazê-lo. Entretanto, no final de 2013, revolvendo papéis e pastas velhas, deparei-me com um envelope empoeirado onde se encontravam as produções dos meus antigos alunos e a pequena fita-cassete com a gravação das apresentações, necessárias para aqui ilustrar de forma honesta e sem acréscimos o que se passou naquela aula em meados de agosto de 2008.

Sobre as motivações que me levaram a arquitetar, propor e conduzir tal atividade, a mais importante, e até hoje presente, era a de demonstrar aos nossos alunos de Teoria da Literatura, visto que a disciplina que ministrava tratava-se de um ramo dessa, que a aplicabilidade dos conhecimentos que adquirimos ao longo de nossa formação vai além da capacitação para o exercício da docência (no caso dos alunos de Ensino Superior) ou da preparação para aprovação em concursos vestibulares e/ou similares (referente aos alunos de Ensino Fundamental e Médio). O intuito agui era conseguir encontrar e oferecer na esfera prática um (e apenas um dentre tantos outros possíveis) exemplo que respondesse a um questionamento corriqueiro acerca de determinados conteúdos, mas que pode fazer tremer nas bases professores das mais diversas áreas: "Qual a utilidade disso para o meu [do aluno] futuro?". Inicio com uma breve contextualização sobre a natureza da disciplina que ministrava, seu público e a preparação para a atividade. Passo, então, ao relato da mesma, seguido da discussão teórica decorrida em sala de aula no durante e após seu cumprimento.

#### 1 - Sobre a atividade e seu início

A disciplina Relações de transtextualidade (60 horas) foi oferecida pelo Departamento de Letras da Universidade Federal do Piauí como cadeira optativa aos(as) graduandos(as) de letras Português, Inglês e Francês. As aulas, por mim ministradas, aconteceram às terças e quintas, das 20 às 22 horas, no segundo semestre de 2008. A proposta da cadeira era apresentar e discutir as cinco relações transtextuais<sup>2</sup> ilustradas por Gerárd Genette em Palimpsestes (1982) através de uma abordagem intersemiótica que tomaria como ponto de partida e/ou retorno o literário, com o objetivo de capacitar os alunos a identificar e distinguir sua presença, assim como perscrutar sua função na estética dos objetos analisados, fossem eles literários ou não. Por se tratar de uma disciplina optativa, aberta a todas as graduações de Letras, o público matriculado era bem heterogêneo no que dizia respeito à fase em que cada um se encontrava em seus respectivos cursos, assim como às razões que os fizeram matricular-se na disciplina. Dos 30 alunos matriculados, apenas 25 permaneceram até o final do semestre, tendo a maioria dos remanescentes em comum o fato de estarem em fases mais avancadas, muitos cursando o último semestre do curso.

Para a atividade, solicitei, com uma aula de antecedência, que os alunos trouxessem para a sala de aula revistas, tesoura, cola e pincéis marcadores. Sem maiores explicações, disse-lhes apenas que os materiais seriam utilizados em uma atividade teórico-prática. Na aula seguinte, materiais em mãos, compareceram 20 alunos, divididos em 4 grupos com 5 membros. Cada grupo recebeu, aleatoriamente, uma cartolina branca de 44 X 61cm, na posição retrato, com um minitexto de aproximadamente 300 caracteres colado no meio, num total de 4 cartolinas. Os textos haviam sido impressos em posição paisagem em folhas sulfite tamanho A4 e não possuíam nenhum elemento

<sup>&</sup>lt;sup>2</sup> Intertextualidade, metatextualidade, paratextualidade, arquitextualidade e hipertextualidade.

paratextual (títulos, subtítulos, figuras, palavras-chave e referências à fonte) que os identificasse.

#### Texto o1

Ela, à beira do lago, observa o horizonte amplo, testemunha de seus passos. Caminha sobre as lembranças guerreiras, inovadoras, sempre alegre como a chegada num porto. Pára, olha em volta, sorri...

Sente-se mãe acolhedora que abraça seus filhos nas sombras de seus jacarandás.

#### Texto 02

As feridas de ontem ainda estão abertas e a dor que está por vir já arde na pele rasgada, crivada. Quantas dores a Humanidade se infligirá? Quantos Cristos terão que renascer? O pensamento se esvai quando escuta o barulho das correntes que se aproximam. Fecha os olhos e anoitece a infância.

## Texto o3

Estava em estado zen quando escutou o grito que vinha do outro lado do muro. Desceu as mãos ainda juntas, ergueu os olhos ao alto e em seguida os baixou. Ergueu-se da posição de lótus, abriu o portão do silêncio e bradou pela liberdade de cultuar a paz.

### Texto 04

De repente a dor parou. Podiam apertar-lhe a língua com alicate, queimar-lhe, algemá-la, dar-lhe excrementos de cão para comer. Estava anestesiada. E nem era ódio... Era um sentimento que amortecia seus sentimentos. Era um descaso pela sua vida que nem ela mais tinha complacência com seu destino.

De posse dos textos, expliquei a primeira parte da atividade, que consistiria na leitura conjunta de cada texto dentro de seu respectivo grupo. Terminada a leitura, os membros dos grupos deveriam debater,

internamente, sobre o que tratava o texto que seu grupo tinha em mãos. Finalizada essa etapa, após chegarem a um denominador comum, os grupos recorreriam às revistas que possuíam em busca de figuras, caixas de texto, títulos, subtítulos, etc. Materiais esses que seriam recortados e colados com o intuito de completar os arredores do corpo textual com base nos elementos que cada grupo julgasse apropriados à transmissão do conteúdo do seu texto. Foi permitido aos grupos o intercâmbio de materiais entre si e incentivado que dessem asas à criatividade. O tempo para preparação da atividade foi de 30 minutos, ficando os 90 minutos restantes reservados para a apresentação e discussão dos "textos dos grupos". O trabalho de recorte e colagem transcorreu sem problemas e a maioria da turma se mostrou bem disposta e compenetrada no cumprimento da atividade.

Textos preparados, solicitei então aos(as) alunos(as) que decidissem uma ordem de apresentação dos grupos e os apresentassem – os textos – à turma, explicando do que julgavam tratar cada um e o porquê de cada elemento escolhido para compor sua versão final do texto<sup>3</sup>.

\_

<sup>&</sup>lt;sup>3</sup> O conteúdo das apresentações foi gravado em fita-cassete, além de compilado em anotações esparsas feitas pelos próprios grupos e por mim durante a atividade. O conteúdo de cada apresentação, seja em citação direta ou em paráfrase, será referenciado pelo ano, 2008.

## 2 - Das apresentações dos grupos

Eis o texto apresentado pelo primeiro grupo:



Figura o1: Texto o1. Arquivo pessoal do autor<sup>4</sup>.

\_

<sup>&</sup>lt;sup>4</sup> Como pode ser observado na figura, foram feitas correções na pontuação da primeira sentença do texto impresso na cartolina a pedido do professor.

Segundo o grupo, o texto foca em um tema bem específico: a maternidade, que se justifica pelo fato da narração em terceira pessoa aludir diretamente à figura materna através do pronome feminino ela (linha o1), somado aos substantivos mãe (linha o6) e filhos (linha o7). Em seguida, os adjetivos utilizados, i.e. guerreiras, inovadoras, alegre (linha 04) e acolhedora (linha 06), de acordo com os alunos, "servem para reforçar uma imagem materna idealizada" (2008)1. Sobre a atmosfera do texto, foi acordado pelo grupo se tratar de uma atmosfera tranquila, especialmente pela presença de elementos naturais na paisagem descrita, i.e. à beira do lago (linha o1), horizonte amplo (linha o2), à sombra dos seus jacarandás (linhas 07-08), que "transmitem uma ideia de paz interior" (2008). Contudo, segundo os alunos, o texto deixa transparecer que a atmosfera atual foi construída a partir de algum esforço anterior, representado por lembranças guerreiras (linha 04), somadas à vontade constante de modificar a realidade, ilustrada pelo adjetivo inovadoras (linha 04), mas encarando o futuro sempre de forma positiva, sempre alegre (linha 04), ciente das surpresas que ele poderá trazer, como a chegada num porto (linha o5). Em resumo, para o grupo, o texto pareceu "tratar de uma mãe orgulhosa e em paz consigo mesma que colhe os frutos de seu esforço e aprecia os resultados de sua dedicação através de seus filhos" (2008).

A escolha dos elementos retirados das revistas é condizente com a proposta do texto, segundo o grupo. No que concerne ao título, *Laços de amor e amizade*, o mesmo "ilustra a relação que se espera entre mães e filhos" (2008), sendo reforçado por outras palavras dispostas ao redor do texto, como os verbos *compartilhar*, *sorrir*, *cuidar*, *aliviar e curar*, *o adjetivo especial*, o grupo nominal *sensações da natureza* e a frase "É como um Conto-de-Fadas". A mensagem é também corroborada pelas figuras utilizadas pelos alunos, a começar pela ideia de família, através de alusões à figura materna, como é o caso das imagens nos cantos

-

<sup>&</sup>lt;sup>1</sup> Notas e gravação da aula de Relações de transtextualidade ministrada pelo professor Cláudio Augusto Carvalho Moura - Centro de Ciências Humanas e Letras — UFPI, 2º semestre de 2008.

superior esquerdo e inferior direito. Na primeira observa-se, conforme relato do grupo:

um pequeno varal com três fotos: primeiro a de uma mulher que parece estar passando por um momento especial, a de uma criança (um menino) com uniforme escolar e por último uma foto que parece ser da mesma mulher junto a um bebê (que pode ser a criança da foto anterior) e uma outra criança, uma menina, que "deve" [grifo nosso] ser a irmã mais velha. Esse varal de fotos remete à família e à passagem do tempo com fotos clássicas e mostra a mulher tanto enquanto ser independente quanto como mãe. Na segunda foto, a de baixo, uma mulher branca abraça uma criança negra e, independente dela ser ou não a sua mãe, a imagem passa a sensação de que o amor não enxerga raça ou outras diferenças. Colocamos logo abaixo dessa figura as palavras cuidar, aliviar e curar para deixar claro que aquela figura se referia, no "nosso" [grifo meu] texto, ao amor sem preconceitos (2008).

Nos cantos superior direito e inferior esquerdo encontram-se fotos que "dizem respeito à liberdade e à natureza. A árvore florida representaria o jacarandá do texto<sup>5</sup> e o homem de terno brincando na gangorra e jogando papéis para o alto representa relaxar e não levar tudo tão a sério" (2008). Como pode ser observado, próximo à figura do homem de terno aparecem a frase "É como um Conto-de-Fadas" e o verbo sonhar, que indicam "os raros momentos em que conseguimos fugir do correcorre e do estresse da vida e como eles são importantes e gratificantes" (2008). Note-se que os alunos passaram a consolidar suas ideias sobre o texto de forma definitiva com a ajuda das figuras, dos vocábulos e das sentenças, ao ponto de se apropriarem do texto, referindo-se ao mesmo como "nosso texto" e de posicionarem figuras e palavras estrategicamente, com o objetivo de transmitir uma mensagem específica. Quanto à participação do restante da turma durante a

<sup>&</sup>lt;sup>5</sup> Trata-se, na verdade, de uma cerejeira.

apresentação do primeiro grupo, observou-se que aqueles que se manifestaram demonstravam concordar, achando que o trabalho ilustrou de forma satisfatória o conteúdo do texto e que "o grupo soube 'vender o seu peixe' muito bem" (2008).

Findada a apresentação do primeiro grupo, que levou cerca de 10 minutos contabilizadas as colocações da turma, passou-se para o segundo grupo, que expôs o texto a seguir:



Figura 02: Texto 02. Arquivo pessoal do autor.

A primeira coisa que se observa no texto do segundo grupo é o seu impacto visual devido ao rompimento dos limites da cartolina pelos títulos, palavras e figuras, assim como as sobreposições entre esses elementos. Outra coisa a ser levada em conta no seu layout é a utilização de pequenas caixas de texto próximas às figuras utilizadas, correlacionando-as, segundo o grupo, com "as várias coisas que o texto parece querer dizer" (2008).

O texto, conforme seus apresentadores, "parece se referir às tristezas e os sofrimentos que não acabam" (2008), ilustrados por expressões como feridas abertas de ontem (linha 01), pele rasgada e crivada (linhas 02 e 03), anoitecer da infância (linha 08), e palavras como dor/dores (linha o2 e o3), Cristos (linha o4) e correntes (linha o7). Parece tratar "de alguém sem esperança que conhece o sofrimento, cujas 'feridas de ontem ainda estão abertas' e que já tem a dor de amanhã como certa" (2008). Nesse contexto, a palavra Cristo, conforme apontado pelos alunos conota sofrimento, mas não um simples sofrimento e sim um sofrimento tal qual a via-crucis, "que funciona como uma imagem muito forte" (2008). A alusão serve para "mostrar que o sofrimento no texto é desumano e tem origem nas ações dos homens, pois são sofrimentos que a própria humanidade se inflige, fazendo do homem vítima dele mesmo" (2008). Prontamente, foi feita por um aluno que assistia a apresentação a relação da fala do grupo com o adágio de Plauto – Homo homini lúpus<sup>6</sup> – popularizado por Thomas Hobbes.

Sobre a disposição e o tipo de figuras e palavras, além das caixas de texto utilizadas, o grupo explicou que o mais importante para eles era "passar uma ideia de sofrimento e de confusão" (2008), o que explica o rompimento das medidas da própria cartolina como "uma forma de mostrar as coisas fora do lugar, como no texto. Mais ou menos como o começo de um delírio" (2008). Questionados sobre o porquê do posicionamento do título, *O valor da tristeza*, no quadrante superior esquerdo, dividindo meio a meio o espaço superior da cartolina com

<sup>&</sup>lt;sup>6</sup> O homem é o lobo do homem.

duas figuras, além do verbo *escolher* como subtítulo, os membros do grupo afirmaram:

Isso não foi pensado conscientemente, mas depois que fizemos aquela disposição achamos que ela estava boa, porque as fotos – um homem sendo arrastado na rua por policiais e uma pessoa sendo retirada por bombeiros de escombros – complementavam o título **como se os dois fossem uma coisa só**. A palavra *escolha* aparece como subtítulo, mas ela é grande, quase do mesmo tamanho do título, e passa a ideia de que o sofrimento é sempre uma escolha, seja de quem opta por sofrer ou de quem opta por fazer os outros sofrerem (2008. Grifos meus).

Na parte inferior da cartolina lê-se, numa caixa de texto, a seguinte mensagem:

"O ditador norte-coreano Kim Jong- II fez seu primeiro teste nuclear e pode desencadear uma corrida armamentista se a ONU não o punir como exemplo".

De acordo com os alunos, a explicação para o seu uso no texto é simples, pois "ilustra o poder que poucos têm de interferir na vida de muitos e como isso pode ser perigoso" (2008) e se, por um acaso, parecer que ela "não se liga diretamente ao texto [é] porque esse texto é bem aberto e geral" (2008). A figura logo abaixo da caixa de texto mostra soldados em campo com bandeiras norte-coreanas em mãos, uma indicação de que figura e caixa de texto, provavelmente, foram retiradas da mesma fonte. Sob ela aparece a palavra terrorismo em letras garrafais, ali posicionada porque "todas essas coisas estão ligadas" (2008).

Ainda do lado inferior esquerdo temos a figura de uma mulher caída ao chão chorando, encostada no que parece ser uma porta de madeira, tendo acima a frase "Onde eu errei?". Essa figura "serve para complementar o que já foi dito, que às vezes o nosso sofrimento é causado por nós mesmos, não pelos outros" (2008). Na parte inferior do

canto oposto se encontra a figura de um coração partido onde se sobrepõe uma caixa de texto que diz:

"Se porrada educasse as pessoas, bandido saia da cadeia santo" LULA, presidente da República, defendendo mais investimentos em comunidades carentes

Acima da caixa a palavra *miséria* foi posicionada, palavras do grupo, "para lembrar que boa parte do sofrimento e das tristezas nascem à partir do momento em que o homem deixa de ter compaixão pelo seu semelhante" (2008), numa alusão clara aos direitos humanos. Por último, quase totalmente fora da borda inferior da cartolina, encontrase o verbo *viver*. Segundo a conclusão do grupo, "(sobre)viver seria o que restaria após o *anoitecer da infância*, em consequência da perda da inocência, dos sonhos destruídos ao longo da vida e da falta de esperança no futuro" (2008). Pode-se notar a expressão *anoitecer da infância* como síntese da atmosfera do texto, apontada pela classe como extremamente obscura e pessimista. A quantidade de informações trazidas pelo grupo tornou sua apresentação a mais longa, tomando cerca de 15 minutos da aula e contando com grande participação dos outros alunos.

Na sequência foi apresentado o texto do terceiro grupo, conforme seque na próxima página:

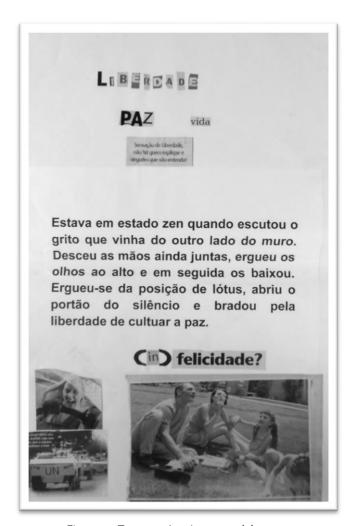


Figura o3: Texto o3. Arquivo pessoal do autor.

É possível, de acordo com o grupo, predizer a partir de expressões e palavras como zen (linha o1), mãos juntas (linha o3), posição de lótus (linha o5) e paz (linha o7) que o texto faz referência à espiritualidade, mais precisamente uma espiritualidade associada ao oriente. Contudo, em síntese, "a mensagem do texto parece estar ligada à luta pelo direito

de exercer uma espiritualidade. Nesse caso à necessidade de se tornar consciente que às vezes se precisa lutar por isso" (2008). O alerta vem "do grito do outro lado do muro, que ilustra como as instituições religiosas tendem a se isolar do mundo" (2008). O grito, para o grupo, foi o responsável por despertar o cidadão, fazendo-o "sair de sua zona de conforto" (2008) e abrir o portão do silêncio (linhas 05 e 06) para bradar pela liberdade de cultuar a paz (linhas 06 e 07). Por meio dos verbos abrir e bradar "fica clara a inconformidade, a disposição para mudar a situação atual". (2008).

O substantivo liberdade (no topo), seguido dos substantivos paz e vida (logo abaixo), fariam as vezes de título do texto: Liberdade, paz e vida, uma vez que, segundo o grupo, apesar da falta da vírgula e da conjunção, o simples posicionamento das palavras na parte superior do texto já seria suficiente para passar a ideia de título, "pois essas coisas estão ligadas e dependem uma da outra" (2008). Para complementar o sentido pretendido, os alunos utilizaram como subtítulo uma caixa de texto onde se lê "Sensação de Liberdade, não há quem explique e ninguém que não entenda!".

Abaixo do texto encontramos uma indagação ambígua, i.e. (in) felicidade?, que os alunos explicaram estar ali para ilustrar que ambas as sensações, "felicidade e infelicidade [,] são apenas lados da mesma moeda. Por isso, ao falarmos de uma, estamos falando, mesmo que indiretamente, da outra" (2008). Uma particularidade que se pôde observar no texto do terceiro grupo foi o fato dos parêntesis de (in) felicidade terem sido feitos à mão, com marcador, pelos próprios membros do grupo, configurando um acréscimo outro que não o previamente solicitado, que pedia que as adições à cartolina fossem feitas a partir do material encontrado nas revistas. Ao acréscimo inusitado o grupo prontamente respondeu que, uma vez

se tratando do **nosso** [deles] **texto**, não havia muita diferença em acrescentar algo que viesse de uma revista ou da cabeça da gente. [Um outro membro complementa] Se bem que quando a gente acrescenta algo por nossa conta, a gente deixa de depender só das revistas que a gente tem aqui. Porque na nossa cabeça tem muito mais coisa que a gente podia usar do quê o que a gente encontrou nessas revistas, porque a gente já sabia o que queria dizer, mas pro texto dizer isso a gente precisou mexer no texto nós mesmos [sic] (2008. Grifos meus).

Seguia logo abaixo de (in) felicidade? a figura de uma família sentada na grama em um momento de descontração. Sua função era ilustrar que "para muitas pessoas a família é o ideal de felicidade, mas [que] isso é relativo, porque muitas vezes nossos problemas podem estar dentro de nossas famílias. Por isso a pergunta está em cima [sic] da foto da família" (2008). À esquerda da figura encontram-se duas menores, uma mulher dirigindo e sorrindo (acima) e um comboio de tanques de guerra (abaixo), que representariam tanto os momentos de felicidades quanto momentos de tensão, assim como nos fariam "lembrar que a felicidade é uma coisa particular. Uns estão felizes enquanto outros não estão e isso muda, porque a vida é assim" (2008). Finalizando a apresentação do grupo, um membro ressaltou que não podia deixar de comparar o trabalho por eles apresentado com os trabalhos dos grupos anteriores e pontuou que se pudessem "ter usado a internet como fonte poderiam ter apresentado um trabalho mais rico, pois a daria pra fazer uma busca exata daquilo que a gente [o grupo] queria passar com o texto" (2008. Grifos meus). Prontamente, toda a turma se manifestou em concordância.

O quarto grupo (o último a se apresentar), tomando a deixa do grupo anterior, inicia sua apresentação justificando os poucos elementos utilizados no seu texto pelas limitações dos materiais em mãos. A seguir a produção do grupo:

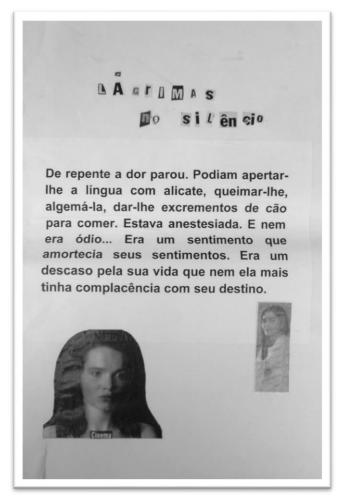


Figura 04: Texto 04. Arquivo pessoal do autor.

A primeira coisa apontada pelo grupo foi o fato da narração em terceira pessoa girar em torno de uma figura feminina – definida pelo adjetivo anestesiada (linha o4) e pelo pronome ela (linha o7) – vítima de tortura. Tortura essa explicitada nas quatro primeiras linhas com "podiam apertar-lhe a língua com alicate, queimar-lhe, algemá-la, dar-lhe

excrementos de cão para comer" (2008). Ainda sobre a situação da personagem, o grupo relatou se tratar de um caso de tortura continuada, pois sua indiferença à dor e a tudo que lhe era infligido denotavam que "aquela situação parecia vir se prolongando há muito e não era novidade para ela. O fato de 'a dor haver parado' não significava o fim de uma sessão de tortura, mas sim que aquilo já havia se tornado corriqueiro" (2008). Para o grupo, o verbo poder (linha 01) pede uma reação em resposta às ações que lhe precedem, reação essa que, por não se manifestar, deixa clara a apatia da personagem, explicitada pelo adjetivo anestesiada (linha 04).

Todo o texto, segundo o grupo, é bastante perturbador, e ilustra "toda a degradação e perda da dignidade da personagem, que aparentava não se sentir mais um ser humano dentro daquela situação. Não havia um amanhã. Não havia nada pra ela além do sofrimento com o qual ela se acostumou" (2008). A realidade da personagem levou o grupo a aludir a um caso recente à época, amplamente noticiado por toda a imprensa mundial. Tratava-se do caso da senadora franco-colombiana Íngrid Betancourt, mantida refém por seis anos pelas Forças Armadas Revolucionárias da Colômbia (FARC) e libertada em julho de 2008. De acordo com o grupo:

O texto podia muito bem ser sobre a Íngrid Betancourt, pois as condições desumanas às quais a personagem é submetida lembram muito o que os jornais dizem que ela passou no cativeiro: tortura, estupros, doenças. Mas o que mais lembra o texto são as imagens da filmagem do rosto, que os jornais mostraram várias vezes, da expressão dela (Íngrid Betancourt) enquanto ainda era prisioneira: uma mulher sem vida, um zumbi. Dá pra ver nas entrevistas que ela dá que mataram alguma coisa dentro dela e que ela nunca vai esquecer tudo aquilo. Vão ser os fantasmas que vão assombrar ela sempre. Que nem a personagem do texto, tiraram a humanidade dela (2008).

O título *Lágrimas no silêncio* ilustraria, conforme o grupo, "toda a atmosfera de dor resignada que o texto apresenta" (2008) e foi montado a partir da colagem de letras, sílabas, ou grupos de letras isolados, retirados de diversas fontes, uma vez que não foi encontrado nas fontes título ou palavras aleatórias que se enquadrassem no que o grupo havia concordado ser o tema do texto. Foram utilizadas também duas figuras de mulheres com expressões tristes. Dentre elas, uma recebeu destaque especial:

Como podem observar, essa foto aqui [foto da direita, abaixo do texto] é da atriz Letícia Sabatella, da [rede] Globo. Escolhemos essa foto porque a atriz já fez várias personagens dramáticas, que sofrem muito e quando olhamos para a imagem dela associamos diretamente ao texto. Como muitas pessoas assistem novelas, imaginamos que a imagem da atriz transmitiria a mensagem que achamos que o texto passa de uma forma mais direta, pois a imagem dela já está ligada a esse tipo de personagem (2008).

Após esse comentário, deu-se por encerrada a apresentação. A despeito da apreensão do último grupo quanto à recepção do seu trabalho, a turma reagiu de forma bastante positiva, concordando com o que foi apresentado e fazendo, ocasionalmente, colocações que corroboravam a linha de interpretação do grupo.

# 3 - Apresentando e discutindo os textos originais

Findadas as apresentações, era chegada a hora de expor e discutir a teoria por trás da atividade. Assim, comecei revelando a origem das fontes com as quais os(as) alunos(as) haviam trabalhado: textos provenientes de 4 dentre 10 postagens disponibilizadas entre os dias 26 de fevereiro e 26 de março de 2008 no *blog Minicontos Cotidianos*<sup>7</sup>, de autoria de Sílvio Vasconcelos. O *blog* em questão trata de uma iniciativa do autor de se apropriar de notícias veiculadas em jornais e reescrevê-

\_

<sup>&</sup>lt;sup>7</sup> http://minicontos.blogspot.com.br

las na forma de minicontos, imbuindo-lhes uma literariedade que aprendemos que o texto jornalístico usualmente não possui.



Fig. 05: Apresentação do blog *Minicontos Cotidianos*. Fonte: http://minicontos.blogspot.com.br

Caso o leitor questione o porquê do uso dessas fontes à margem do cânone, explico que a natureza da disciplina, que parte do estudo do texto independente de questões quais a literariedade, é bastante aberta, servindo não apenas para a literatura, mas também para áreas como a linguística, publicidade, jornalismo, entre outras tantas. Ademais, fosse qual fosse a fonte utilizada, todo o conteúdo apreendido em sala poderia ser utilizado para análise de obras literárias, fossem elas canônicas ou não. Logo, restringir os materiais de análise utilizados apenas a produções reconhecidas e aclamadas seria restringir o leque de possibilidades de aplicação dos conteúdos trabalhados. Trabalhar com outras fontes, ao contrário, serviria para relembrar aos(as) alunos que os estudos literários são também aplicáveis a outras áreas e que através do seu componente basilar, o texto, todas estão ligadas. Não posso negar também o fato de que o blog foi um achado, pois suas potencialidades didáticas atendiam exatamente ao que eu tinha em mente para aquela atividade, satisfazendo suas necessidades mais do que qualquer manual de Teoria da Literatura ao meu alcance à época.

Após a revelação da fonte era a hora de mostrar aos alunos os originais dos textos sobre os quais eles haviam trabalhado. Tal qual o conteúdo das apresentações dos alunos, é importante que o leitor deste texto possa visualizar exatamente do que falo. Por isso, trago aqui não apenas a discussão que seguiu após a apresentação de cada texto original, mas o texto em si, conforme retirado do *blog*, para que o leitor possa também formular suas próprias conclusões.



Figura o6: Versão original do Texto o1. Fonte: http://minicontos.blogspot.com.br/2008/03/236-anos-em-360-graus.html

Como pode ser agora observado, o primeiro texto tratava, inicialmente, de um texto em comemoração do 236º aniversário da cidade de Porto Alegre -RS. Publicado no dia 26 de março de 2008, o texto "236 anos em 360 graus" era uma releitura de uma matéria veiculada pelo jornal Zero Hora à mesma data, originalmente intitulada "Porto Alegre de todos os ângulos". Pedi então à turma que dissesse como foi possível retirar todas essas informações desse "novo" texto, ao que pronto me responderam que a citação da fonte logo abaixo da figura, seguida do título original e dos hiper-marcadores, deixavam evidente a sua origem. Ao comparar o texto do grupo – "Laços de amor e amizade" – ao texto do blog ficou claro que o conteúdo do corpo textual, isoladamente, era condizente com ambas as versões. As adições que circundam o texto apresentado pelo grupo não alteraram o seu interior, que permaneceu incólume, sendo o corpo textual per se suficiente para fazer com que o conteúdo mantivesse uma "aura original", vez que uma relação entre os caracteres mãe e cidade é completamente plausível. Contudo, essa relação só ficou clara após a apresentação do texto secundário, desvelando a estrutura palimpséstica que se havia estabelecido entre os textos, incluindo aí o real hipotexto, a matéria de jornal em si, à qual não tivemos acesso.

Ao comparar ambos os textos, além da óbvia riqueza visual do texto dos alunos em relação ao original, é possível observar que os elementos por eles incorporados ao seu texto suscitam vários aspectos diretamente relacionados ao tema da maternidade, enquanto no texto original apenas o título remete de forma direta ao conteúdo: o aniversáro de Porto Alegre. Em detrimento ao tom mais explícito que o texto do grupo adquiriu com as adições, o texto original, justamente por não têlo feito, cercou-se de um certo ar poético. Portanto, é plausível afirmar que os elementos acrescentados foram responsáveis por concederem ares distintos à mesma narrativa. Como resultado, apresentaram-se, no confronto entre os dois textos, duas narrativas aparentemente independentes que compartilham, entretanto, o mesmo corpo textual, que permaneceu inalterado.

O segundo texto, nomeado pelo seu grupo "O valor da tristeza", chamava-se, enquanto miniconto, "Em cada manhã, uma nova escuridão", inspirado no texto da matéria intitulada "Achava que estava educando, diz suspeita de tortura", veiculada em 24 de março de 2008 também pelo jornal *Zero Hora*.



Fig.07: Versão original do Texto 02. Fonte: http://minicontos.blogspot.com.br/2008/03/em-cada-manh-nova-escurido.html

Após uma olhada na foto e nos hiper-marcadores (escravidão, infância) os alunos chegaram à conclusão de que o texto tratava de uma criança

que vivia em regime de escravidão e sofria torturas. De posse dessa informação, a turma começou a inferir que as correntes (linha o6) citadas no texto poderiam muito bem ser correntes reais que serviam para prender a criança, ao invés das correntes metafóricas imaginadas pelo grupo enquanto trabalhava apenas com o corpo do texto destituído de qualquer outro elemento referencial. Apesar disso, se o leitor voltar ao relato da apresentação do segundo grupo poderá constatar que as conclusões que os alunos chegaram acerca do conteúdo do corpo textual aproximam-se claramente da mensagem veiculada pelo miniconto. Relembro em especial o adágio trazido durante a apresentação: "o homem é o lobo do homem", aplicável a ambas as situações, assim como toda e qualquer situação de abuso. Mais uma vez um texto funciona como palimpsesto do outro, descoberto pelos elementos circundantes do texto original. Contudo, esses elementos reforçam não a sua generalidade, os seus elementos comuns, mas sim as especificidades de cada.

Entre o texto do blog e o texto apresentado pelo segundo grupo, a turma foi unânime em expressar sua preferência pelo texto apresentado pelos colegas. Segundo os alunos os elementos utilizados por seus companheiros de classe enriqueceram o texto, pois aumentaram muito seu potencial de referenciação em comparação ao primeiro. Para os alunos, a discussão fomentada pelo trabalho do grupo foi muito mais profícua do que seria uma discussão feita a partir do miniconto, pois os elementos incorporados ao corpo textual pelo grupo ofereceram ao mesmo uma profundidade maior que aquela perceptível em sua fonte.

Na sequência foi apresentado o miniconto trabalhado pelo terceiro grupo. Intitulado "Entre o silêncio e o grito", o texto teve como inspiração a matéria "O Dalai Lama: pronto para conversar com líderes chineses" do jornal italiano *Corrieri della Sera*, edição de 20 de março de 2008.

\_

<sup>&</sup>lt;sup>8</sup> No original: *Il Dalai Lama: pronto a parlare coi leader cinesi*.



Fig. 08: Versão original do Texto 03. Fonte: http://minicontos.blogspot.com.br/2008/03/entre-o-silncio-e-o-grito.html

Do corpo textual ao qual o grupo teve acesso resultou o texto "Liberdade, paz, vida". Relembrando sua apresentação podemos constatar, com uma breve olhada no título, figura e hiper-marcadores do miniconto, que eles acertaram ao predizerem, pelos elementos

disponíveis dentro do texto ao qual tiveram acesso, que o texto fazia referência a uma forma de espiritualidade oriental. O argumento holístico apresentado pelo grupo de que as palavras que formavam o título, a despeito do não uso da vírgula para separá-las, estavam ligadas umas às outras como essencialmente complementares e interdependentes entre si, denotam o conhecimento de conceitos da cultura oriental que, embora não completamente compreendidos pelo ocidente, encontram-se massivamente difundidos através dos meios de comunicação. A relação feita pelo grupo, contudo, só se tornou explícita durante sua apresentação e através da fala de seus membros, o que levou os alunos a discutirem o quão importante é o conhecimento de mundo para que se possa enxergar relações que muitas vezes se encontram veladas, mas nem por isso menos presentes, e o quanto elas – essas relações – podem vir a enriquecer a interpretação dos textos.

A turma apontou quão mais claro tornou-se, ao comparar o texto original e sua versão, perceber o quanto a figura do monge, com uma faixa na cabeça onde se lia Free Tibet sendo levado aos gritos por um policial, que ilustra o miniconto é, em si, contrastante, se considerado o descompasso dessa ação em relação à ideia de paz e tranquilidade que a imagem monástica usualmente remete. Há no texto do grupo também esse tipo de contraste, mas ele se mostra de duas formas: no plano pictórico temos fotos com cenas de felicidade e a foto de tanques de querra, no plano textual temos as palavras felicidade X infelicidade, representadas pela montagem (in)felicidade?, disposta acima da figura maior na parte debaixo do texto. Foi possível aos alunos perceber que o contraste apresentado no original requer o conhecimento de um contexto mais específico e por isso pode se dar de forma mais sutil, dependendo do tipo de receptor. Já o contraste apresentado pelos elementos trazidos pelo grupo, além de se mostrar de forma dupla em cadeias sígnicas diferentes, não exigia nenhum conhecimento específico, pois ilustrava de modo mais acessível um conflito inerente à condição humana. Também nessa apresentação a turma concordou que a versão do grupo era mais propícia à suscitar discussões que o texto original devido à maior quantidade de elementos distribuídos nos arredores do corpo textual.

O último texto era também baseado em uma matéria do jornal *Zero Hora*. Publicada em 19 de março de 2008, a matéria se intitulava "Menina torturada em GO encontra pais biológicos" e dela derivou o miniconto "Dores", publicado à mesma data.



Fig.og: Versão original do Texto o4. Fonte: http://minicontos.bloqspot.com.br/2008/03/dores.html

À exceção do texto tratar de uma menina, consoante o título da matéria original, também a interpretação do último grupo manteve-se próxima à do texto de origem. Conforme os alunos, a particularidade desse texto em relação aos outros era o fato de sua escrita explicitar o seu conteúdo de forma mais direta que os anteriores, não deixando muita margem para outras interpretações. Comparando o texto do grupo ao miniconto foi apontado primeiramente que o título "Lágrimas no silêncio" é mais condizente com a mensagem do texto do que o título que aparece no blog, "Dores". O primeiro seria mais específico ao passo que o segundo seria muito genérico. Em contrapartida, comparadas as figuras utilizadas em ambos, a figura do miniconto remete de forma mais direta ao conteúdo se consideradas as explicações que o grupo deu sobre o porquê do uso da figura da atriz – Letícia Sabatella – em seu texto. Embora ambas ilustrem, superficialmente, mulheres que passa(ra)m por situações de abuso, a figura da menina abraçada ao pai explicita o conteúdo que se pretende transmitir de maneira a permitir uma assimilação mais fácil desse, pois ilustra literalmente o conteúdo do título. Já a figura da atriz, para que fosse compreendida a dimensão do sofrimento ao qual o grupo fazia referência, demandaria todo um conhecimento prévio sobre o tipo de personagens que ela interpretava, o que exigiria ao leitor ir bem além do texto.

Todas as observações sobre uma maior especificidade do último texto serviram para validar o argumento pré-apresentação do grupo sobre a dificuldade em encontrar elementos que ilustrassem de modo satisfatório o seu conteúdo. Fora o título, que pôde ser montado a partir do recorte de letras e sílabas (mas nem por isso exigiu menos trabalho), os outros elementos se mostraram mais difíceis de serem encontrados ao acaso, apenas utilizando as revistas disponíveis. Dificuldade que, conforme observado por um dos membros do grupo, seria contornada se a atividade permitisse a utilização de alguma ferramenta de busca por imagens na internet. Retomou-se mais uma vez a semelhança do conteúdo do texto com o episódio do sequestro de Íngrid Betancourt, relatado durante a apresentação do grupo, para exemplificar que os resultados de uma busca por seu nome na *internet* já forneceriam

material suficiente para um trabalho mais rico. Isso para citar apenas uma dentre tantas possibilidades de material que poderiam ser encontrados e utilizados com uma busca rápida através de outras palavras ou expressões-chaves ligadas a um contexto de violência contra a mulher.

## 4 - Explicitando o implícito: da paratextulidade

Conforme previamente dito e reforçado, o objetivo desta atividade era fazer com que os alunos pudessem reconhecer, através da prática, como funciona o mecanismo da paratextualidade, para a partir daí entender a extensão de sua aplicabilidade. Ao invés de uma explanação teórica inicial acompanhada de exemplos, precedida ou não pela leitura de textos acadêmicos, seguida de alguma forma de checagem da apreensão do conteúdo, optei por tentar outro caminho. Apostei no potencial do material e da atividade para despertar nos alunos um olhar crítico por sobre o conteúdo que desejava abordar. Partindo do preceito que toda teoria tem seu embrião na observação e formulação de hipóteses acerca de determinado objeto, lancei mão do objeto simplesmente, tal qual ele era, antes de qualquer enquadramento teórico, qualquer taxonomia, e dei um crédito tanto aos instintos dos alunos quanto aos conhecimentos que cada um trazia consigo. Baseeime no preceito simples de que o mais importante em nossa labuta, enquanto estudiosos de Literatura, é o texto literário, e que as teorias são apenas instrumentos, muitas vezes contestáveis, mas que partem dele, por nós usadas na tentativa de entendê-los, ou mesmo para alguns, domá-los. Caso fosse o contrário, nasceriam das teorias as obras e, ao viverem essas em função daquelas, não teríamos mais uma Teoria da Literatura, mas sim uma Literatura da Teoria.

Infelizmente esse paradoxo não está tão aquém de nossa realidade e no trabalho estritamente focado na teoria em detrimento do texto, o estudo da Literatura corre o risco de se tornar uma coisa enfadonha e distante. Como resultado, não só a função da Teoria da Literatura passa a ser contestada, como também a nossa relevância profissional, essa

duplamente, uma vez que além de ensinarmos algo de natureza "questionável", falhamos no que concerne à formação dos indivíduos pelos quais somos responsáveis em nosso ofício, pois esquecemos nós mesmos como mostrar-lhes o quão útil pode ser o estudo da Literatura e quanto prazer o seu objeto pode nos proporcionar. No caso dos meus alunos, era importante fazer com que eles paratextualidade, pois só assim poderiam entender efetivamente o seu mecanismo e dar-lhe a devida importância. Não bastava que aquele conteúdo fosse absorvido enquanto pré-requisito para o cumprimento da disciplina. Importava que eles entendessem que poderiam fazer uso daquilo em suas vidas.

Para apresentar formalmente a paratextualidade, lancei aos alunos a proposta de que lessem um texto curto, de minha autoria, parte do segundo capítulo de minha dissertação, que encontrava-se à época em vias de escrita e viria a ser publicada depois como livro . Tratando justamente da transtextualidade em Genette, o excerto, que aqui incluo, resume o que vem a ser a paratextualidade. Após a leitura do texto a turma deveria rediscutir as produções da atividade, agora não apenas instintivamente, mas cientes de toda a teoria por trás do que eles já haviam discutido e que agora lhes era, finalmente, apresentada. Seque uma versão com pequenas modificações do texto entreque aos alunos:

Segunda das relações transtextuais explorada em *Palimpsestes*, precedida pela intertextualidade, define-se como paratextualidade as relações que o texto mantém com os seus paratextos, definidos por Genette como

> título, subtítulo, intertítulo, prefácios, posfácios, advertências, prólogos, etc.; notas marginais, de rodapé, de fim de texto; epígrafes; ilustrações; errata, orelha, capa e tantos outros tipos de sinais acessórios, autógrafos ou alógrafos, que fornecem ao texto um aparato (variável) e por vezes um comentário, oficial ou oficioso, do qual o leitor, o mais purista e o menos vocacionado à erudição externa, nem

sempre pode dispor tão facilmente como desejaria e pretende.(2006,p.08-09)

O paratexto deve ser compreendido como toda e qualquer informação que se faz visível, mas que não está inserida no corpus do texto, complementando e muitas vezes conduzindo o sentido do mesmo. Esse tipo de relação não se faz presente apenas em literatura, mas pode ser observado também no texto eletrônico, por meio dos exemplos de paratextos já citados.

Além de servirem como informação adicional, por complementarem o texto, os paratextos acabam por desempenhar o papel de links ao conduzirem o leitor a um determinado caminho que pode vir a suplementar, ou não, a sua leitura. Mesmo quando não configurar um link formal, um paratexto pode remeter, em um plano mental, a alguma outra realidade que se relacione de alguma forma com o texto, reforçada pela presença da intertextualidade. Funciona, desta forma, como uma espécie de janela para um caminho que se faz apenas pelo plano mental, mas que não deve ser menosprezado, por ser parte da construção de significados do texto pelo seu receptor.

Se observadas essas características, os paratextos passam a ser ferramentas importantes para a captação e interpretação da mensagem tanto em documentos impressos quanto digitais, visto que uma alteração no seu aparato - nesse caso incluímos também a exclusão - pode mudar completamente o sentido de um texto. A partir do momento que essas informações passam a interagir com o texto escrito para sua melhor compreensão, teremos então uma leitura semiótica, pois as cadeias significativas não se limitarão apenas ao signo escrito. Nessa perspectiva os paratextos podem, de fato, ser considerados parte integrante do texto, deixando de ser apenas informação extra, como pode ficar muitas vezes subentendido em uma análise mais superficial dos mesmos (MOURA, 2011, p. 62-3).

.....

Somadas a atividade, apresentações e discussões à leitura do texto, os alunos estavam aptos a formular juízos e conjecturar sobre a paratextualidade de forma menos instintiva e mais organizada. Essa nova, e última, etapa da discussão iniciou com uma reflexão acerca da importância das relações paratextuais.

Uma vez reconhecido na prática o potencial dos paratextos, foi possível ter consciência de que seu uso influencia diretamente na mensagem do texto como um todo. A ele – o texto – associados, os paratextos passam a ser mais que elementos incorporados, tornam-se parte do texto, podendo ser tão importantes quanto, ou até mais, que o corpo textual na transmissão da mensagem. Os textos preparados pelos alunos mostraram que a mensagem do corpo textual poderá ser reforçada, reformulada, enfraquecida, tornar-se geral ou mais específica com a adição de elementos e que, cientes do potencial dessa ferramenta, era possível se darem conta do quão poderosa ela efetivamente era. A publicidade foi citada como exemplo de uso consciente da paratextualidade, em especial as antigas campanhas de cigarros e bebidas, onde a mensagem (que é antes de tudo um texto) vem sempre acompanhada de outros signos e mensagens teoricamente correlatas que reforçam positivamente o consumo desse tipo de produto. No caso dos textos produzidos pelos alunos, o resultado foram quatro textos que diferiam dos seus originais, tanto que os últimos (os originais) só poderiam ser classificados como tais devido à informação prévia de que eles o eram. Do contrário, os textos produzidos pelos alunos poderiam passar-se muito bem por textos autênticos. Os índices paratextuais serviram, portanto, para caracterizar cada texto como único.

Das relações entre os textos produzidos em sala e seus originais, permaneceu como elo o aspecto mais amplo de cada tema trazido no corpo textual, captado pelos alunos. A existência desse elo serve para lembrar que no conjunto textual formado pelo corpo do texto e paratextos cada elemento tem sua importância. A mensagem, que nesse caso encontrou sua especificidade nos paratextos, tem como condutor original o corpo textual, pois foi a partir dele que se fez o arranjo paratextual para da soma de ambos dar-se o resultado final. Extirpados os paratextos, a mensagem continuaria por lá, embora sua apreensão possa ser dificultada, pois no caso dos textos trabalhados havia a pretensão de uma linguagem literária, que não tem compromisso de informar determinado conteúdo de forma clara e objetiva. Os textos que inspiraram os minicontos do blog, entretanto,

tinham por função reportar de forma clara os fatos aos quais faziam referência, pois eram textos jornalísticos. Mesmo possuindo, eles também, o mínimo de elementos paratextuais que fosse (i.e. seus títulos), tais elementos teriam como função informar de forma clara e direta o teor do corpo textual. Se fossem usados esses textos ao invés de seus derivados em sala de aula a atividade não teria tido a fluência que teve, pois, cientes do que cada texto tratava exatamente, os alunos não se sentiriam tão livres e teriam dificuldades em encontrar elementos que satisfizessem um texto com demandas tão específicas. Foi essa, em outras palavras, a justificativa dada pelo último grupo ao final de sua apresentação, tanto que foi reconhecido pela sala que o último texto era menos geral que os três anteriores.

Outro aspecto também trazido para a discussão foi o suporte no qual cada um se apresentou. Os textos entregues aos alunos eram textos impressos, logo seu suporte era físico, enquanto a fonte de onde eles foram retirados, o blog, era virtual. Embora todo texto seja, em essência, virtual, uma vez que o suporte faça, primordialmente, vezes de veículo através do qual ele – o texto – se manifesta, esse mesmo suporte influencia diretamente tanto o seu processo de criação quanto sua recepção, logo, a maneira como vemos o texto mantém com ele uma relação estreita. A atividade de construção de textos a partir da adição de paratextos realizada em sala de aula seria um exercício primitivo da construção de textos na internet, pois sua proposta seria mais facilmente executada em um ambiente de consulta e acesso menos restrito às fontes que lhes parecessem apropriadas. Uma vez consideradas as restrições físicas da atividade (número limitado de materiais em mãos no momento e, por consequência, de conteúdos), aspecto apontado pelos dois últimos grupos, em especial o último, a leitura dos dois tipos de texto também procederia de forma diferente pelas mesmas razões.

Através do texto eletrônico, os paratextos têm a capacidade, como pudemos observar nos minicontos, de explicitar certas referências, seja através dos *hiperlinks* disponibilizados ao final de cada conto, que

conduzem o leitor a textos enquadrados em uma categoria comum, seja através dos títulos originais das notícias, transformados em links que os levariam direto às suas fontes, o que poderia complementar a leitura dos minicontos. Todas essas relações, em se tratando dos textos impressos trabalhados pelos grupos, também são possíveis, com a diferença de se darem em um plano mais sutil, onde as ligações e os elos aconteceram dentro da cabeça do leitor. Nesse caso, o texto eletrônico nos leva diretamente, de modo fluido, a caminhos para os quais o texto impresso exigiria um esforço maior. Apesar de sua aparente facilidade, o texto eletrônico nos abre a porta para outros questionamentos até então não atentados, vez que estamos acostumados com o modus operandi impresso. A função dos paratextos em ambos permanece a mesma. O que difere é a forma como o leitor interage tanto com o texto como com o seu suporte. Difere o grau de explicitação da relação que se faz por trás de cada paratexto, mas uma coisa permanece certa, conforme pôde ser concluído em sala: esses índices, embora muitas vezes subestimados e relegados a um segundo plano por um olhar destreinado, sempre querem dizer algo, e o dizem. Seja no plano da Literatura ou em qualquer outro momento de nossas vidas.

#### **REFERÊNCIAS**

GENETTE, Gerárd. *Palimpsestes*: la littérature au second degré. Paris: Seuil, 1982.

\_\_\_\_\_\_. *Paratexts*: Thresholds of Interpretation (Literature, Culture, Theory). Cambridge: Cambridge University Press, 1997.

MOURA, Cláudio A. C. A abordagem genettiana. In: *Na trilha do hipertexto*: Italo Calvino e As cidades invisíveis. Teresina: EDUFPI, 2011. p.59-71.

VASCONCELOS, Sílvio. *Minicontos cotidianos* [blog]. 2008. Disponível em <a href="http://minicontos.blogspot.com.br/">http://minicontos.blogspot.com.br/</a>>. Acesso em 15 jan 2014.

# Leitura digital em tela: a hipermodalidade como forma de significação textual

Gisella Meneguelli de Sousa<sup>1</sup>

# Cenografias prévias

O presente trabalho tem como objetivo principal analisar a leitura e a escrita de textos produzidos e disponibilizados em meio digital e a emergência de práticas de letramento digital, visto que o ciberespaço já foi institucionalizado por seus usuários como espaço de produção, difusão e leitura de textos das mais diversas esferas do conhecimento. Nesta análise, refletirei sobre algumas estratégias utilizadas pela professora/tutora da disciplina Geografia II do curso a distância de Pedagogia de uma universidade pública brasileira.

É inegável que devem anteceder o letramento digital práticas pedagógicas de alfabetização (isto é, possibilitar a um indivíduo adquirir a tecnologia da escrita pelo reconhecimento dos sinais gráficos da sua língua natural) e de letramento, que é habilitar o indivíduo a ler e escrever textos, reflexiva e criticamente, em sua língua, nos mais diversos gêneros textuais. Barton e Hamilton (1998) dizem que um indivíduo é plenamente letrado quanto é capaz de ir além dos limites da língua e de fazer relações interdiscursivas, vinculando-as à sua realidade histórica, social e política. Não vou me estender nesse ponto, já que existe uma vasta literatura sobre o tema (GRAFF, 1995; KLEIMAN, 1995; SOARES, 1998; ROJO, 2009). Também seria legítimo aventar aqui a

<sup>&</sup>lt;sup>1</sup> Universidade Federal de Juiz de Fora (UFJF). Contato: gisellameneguelli@yahoo.com.br

importância do letramento visual como o fazem Kress (1998, 2005, 2006), Oliveira (2006) e Dondis (2007), que vêm requerendo uma nova teoria que explique adequadamente a linguagem visual e a sua integração com outros modos, seja na escrita, seja na leitura.

Entretanto, vou procurar, primeiramente, situar a teoria cognitiva de leitura *bottom-up* e *top-down* em estudos sobre letramentos, incluídos estudos sobre letramento digital, em relação a teorias que analisam as possibilidades enunciativas advindas da hipermodalidade.

Esta investigação mostra-se relevante porque propõe repensar práticas pedagógicas de leitura e escrita que priorizam o aspecto hipermodal e/ou multimodal de textos, incorporando análises sobre as especificidades das várias linguagens utilizadas na produção de textos, bem como a integração entre elas na geração de múltiplas possibilidades de (re)construir significados.

#### 1 - Formas de processamento de leitura

Há dois tipos básicos de processamento da informação utilizados por leitores em suas práticas de leitura: os processamentos *top-down*, ou descendente, e *bottom-up*, ou ascendente. O processamento informacional descendente (*top-down*), segundo Kato (1995), "é uma abordagem não-linear, que faz uso intensivo e dedutivo de informações não-visuais e cuja direção é da macro para a microestrutura e da função para a forma". Já no processamento ascendente (*bottom-up*), ainda segundo a autora, o leitor faz "uso linear e indutivo das informações visuais, linguísticas, e sua abordagem é composicional, isto é, constrói o significado através da análise e síntese do significado das partes". Várias correntes linguísticas têm privilegiado tanto uma quanto outra abordagem de processamento de informação para descrever tipos de leitores e as estratégias utilizadas por eles em atos de leitura.

O tipo de leitor *top-down* fundamenta-se em um modelo psicolinguístico centrado essencialmente no uso de estratégias de

leitura que resultam de suas experiências prévias adquiridas ao longo da vida, as quais ele utiliza para fazer previsões sobre um texto, ao invés de se centrar nas informações efetivamente constantes no texto. Trata-se de uma estratégia de leitura ancorada mais nas experiências do indivíduo do que nas expectativas do leitor sobre o texto. Já o leitor do tipo bottom-up baseia-se em um modelo de decodificação de informações centrado no texto. A leitura empreendida por esse leitor parte do específico para o geral, ou seja, do nível das palavras ao nível das sentenças, dando pouca margem às atividades cognitivas inferenciais. Logo, a leitura realizada por esse tipo de leitor situa-se em um nível muito superficial, uma vez que a sua capacidade de fazer inferências, estabelecer relações e sintetizar ideias é bastante limitada.

De acordo com Kato, o leitor maduro seria aquele que estabelece de forma consciente o *interactive model*, ou seja, é aquele que

[...] usa, de forma adequada e no momento apropriado, os dois processos complementarmente. É o leitor para quem a escolha desses processos é já uma estratégia metacognitiva, isto é, é o leitor que tem um controle consciente e ativo de seu comportamento. (KATO, 1995, p. 51).

O interactive model é uma interseção cognitiva que situa o leitor em relação ao texto e às suas representações sociais. A leitura é uma prática social, ou seja, a atividade de leitura é uma prática sociocomunicativa situada, na qual a construção de saberes se dá em eventos que envolvem interações com objetivos de leitura específicos, levando em conta a experiência de saberes e crenças internalizadas pelo leitor, bem como a sua capacidade inferencial e interdiscursiva, que é ativada no e pelo texto.

## 2 - Ressignificando práticas sociais de escrita e leitura

Com a popularização dos computadores pessoais e facilitação do acesso à *internet*, novas práticas sociais passaram a ser significadas. Antes da

combinação computador/internet, as pessoas não conversavam via chat, não postavam pensamentos e acontecimentos diários em redes sociais - só para mencionar alguns exemplos. Com o surgimento das mídias móveis, como os smartphones e os tablets, as pessoas passaram a ter em mãos tecnologias informacionais e comunicacionais de forma muito rápida e direta.

Considero importante conceituar práticas sociais, já que será um termo recorrente neste trabalho. Citarei Xavier para apresentar esse conceito:

As *Práticas Sociais* são as formas culturais pelas quais os indivíduos organizam, administram e realizam suas ações e atitudes esperadas em cada um dos diversos Eventos de Letramento existentes na sociedade. Essas ações são, ao longo do tempo, construídas conjuntamente pelos cidadãos comuns e algumas delas passam a ser ritualizadas e oficializadas, posteriormente, pelas instituições que as retomam e exigem que os indivíduos as utilizem em momentos específicos da vida social. Os indivíduos, quando são expostos sociocomunicativamente a esses eventos, tendem a sedimentar usos de formas de fala ou escrita e, assim, os absorvem com certa naturalidade (s/d, p. 6).

As práticas sociais são mediadas pela linguagem, que se materializa em gêneros textuais (orais, escritos e digitais). Nas atividades de letramento, os gêneros assumem um papel essencial no processo de aprendizagem de habilidades letradas, tanto na instância de produção quanto na de recepção.

Naturalmente, a mediação das tecnologias de informação e comunicação (TIC) nos processos de comunicação vem acarretando várias transformações na linguagem expressa na tela do computador e, portanto, novas práticas sociais. Snyder (1997) acredita que o microcomputador propicia um espaço novo de escrita que, consequentemente, impacta as formas pelas quais chegamos a conhecer e organizar os nossos pensamentos.

Se estamos tendo novas experiências cognitivas e novas formas de mediar nossa relação com o conhecimento, a Educação e a Linguística precisam contemplar essa nova realidade que se apresenta aos estudantes.

Visto que um número relevante de práticas sociais de sujeitos em idade escolar e, também, de adultos letrados, ocorre na tela do computador, é importante a indagação sobre quais estratégias de leitura esses leitores demandam e que estratégias podem potencializar a leitura e a compreensão de textos que circulam no ciberespaço. É fundamental que as teorias pedagógicas de ensino de leitura tratem da relação entre o material verbal e o não verbal, como também da linguagem utilizada em textos de caráter multimodal, uma vez que as tecnologias digitais de produção de textos já se consolidaram, permitindo a exploração da expressividade de várias modalidades de linguagem (verbal, visual, sonora) e a sua consequente integração em links digitais — os hipertextos. Em vistas disso, o ensino de leitura, contemporaneamente, não pode deixar de privilegiar os recursos hipermodais e multimodais como integrantes do texto e coparticipantes nos processos de significação textual.

As práticas de letramento escolares, conforme Rojo (2009), dentre outros objetivos, devem possibilitar os letramentos multissemióticos, os quais são "exigidos pelos contemporâneos, ampliando a noção de letramentos para o campo da imagem, da música, das outras semioses que não somente a escrita". Diante da gama de textos disponíveis na tela do computador, o letramento "tradicional", que se ocupa de práticas sociais reguladas principalmente pela escrita, não é mais suficiente para descrever os fenômenos linguageiros originados da integração entre outros modos semióticos.

A institucionalização dos letramentos multissemióticos como práticas de letramento implementa os *letramentos críticos e protagonistas*, cujo objetivo é dar tratamento ético aos discursos situando-os social e contextualmente, além de criar ambientes de problematização dos

discursos hegemônicos e de valorizar a diversidade da polifonia discursiva.

Para isso se tornam tão importantes hoje as maneiras de incrementar, na escola e fora dela, os letramentos críticos, capazes de lidar com os textos e discursos naturalizados, neutralizados, de maneira a perceber seus valores, suas intenções, suas estratégias, seus efeitos de sentido. Assim, o texto já não pode mais ser visto fora da abrangência dos discursos, das ideologias e das significações, como tanto a escola quanto as teorias se habituaram a fazer. (ROJO, 2009, p. 112).

A compreensão de um texto não pode ser considerada uma atividade individual, isolada da sociedade na qual o sujeito vive, uma vez que compreender um texto é uma atividade de interação e não somente uma atividade cognitiva e linguística. A compreensão é uma atividade colaborativa entre autor-texto-leitor ou falante-texto-ouvinte. Logo, o sentido não está no autor, ou no texto, ou no leitor, mas na interação entre esses elementos constituintes de sentidos.

A partir do momento em que a tecnologia digital passou a integrar um conjunto de práticas cotidianas na sociedade, teorias sobre letramento digital emergiram na Educação e na Linguística. Segundo Marcuschi (2008), a língua é um sistema simbólico polissêmico, mas que não tem uma semântica pronta nem autonomia de significado. A exploração dos significados que emanam da multimodalidade textual é um recurso para ampliar a significação textual, uma vez que a integração de diferentes semioses produz novos tipos de textos e de textualidades, além de novas possibilidades de significação.

Um aspecto a se considerar nos estudos de letramento digital, consoante Landow (1997), é o contexto hipermídia, que gerou a necessidade de entender melhor o efeito da organização hipertextual na escrita e na leitura. É necessário entender como viabilizar interações significativas com material digital, não mais apresentando o texto de

forma sequencial (como nos textos orais ou escritos), mas sim através de um conjunto de lexias – segmentos de informação (verbal, visual, sonoro) com sentido completo –, relacionadas entre si através de links digitais, de forma a possibilitar trajetórias variadas de leitura.

A hipermídia é uma forma de organização textual na qual o hipertexto é um agregador de múltiplas semioses. É importante distinguir os conceitos de hipermídia e de hipertexto para a compreensão de como o ciberespaço se organiza.

A linguagem própria do ciberespaço é a hipermídia, que é "a integração sem suturas de dados, textos, imagens de todas as espécies e sons dentro de um único ambiente de informação digital" (SANTAELLA, 2004, p. 48). A hipermídia integra o verbal, o visual (imagens fixas, estáticas, vídeos) e o sonoro em um todo complexo. Portanto, as linguagens do ciberespaço são hipermidiáticas. Já Landow (1997) entende a hipermídia como uma extensão da noção de hipertexto, incluindo informações visuais, som, animação e outros formatos de dados. Xavier (2002) caminha na mesma direção ao afirmar que o hipertexto é formado pela bricolagem de várias linguagens (som, imagem e escrita), que são chamadas também de hipermídia. Este último considera que a hipermodalidade (hipertexto constituído de multimodalidades) é "uma tecnologia enunciativa que viabiliza a emergência de uma nova forma de acessar, produzir e interpretar informações de maneira multissensorial, que se constitui no modo de enunciação digital" (XAVIER, 2002, p. 100).

Se as práticas comunicativas cotidianas passam a ser recorrentes em um outro meio - o digital - é natural que a linguagem tenha outros recursos expressivos explorados. A tecnologia digital favorece a integração, em um único meio, de um conjunto de semioses. Essa combinação "centralizada" de variadas linguagens em um único meio pode propiciar a ampliação de leituras plurissignificativas.

Em vistas disso, diferentes estratégias de leitura são demandadas devido a novas formas de textualidade. Segundo Landow (1997), o hipertexto é um texto composto de blocos de palavras ligados eletronicamente por vários caminhos - denominados de link, nó ou rede. O hipertexto pode se organizar internamente de forma axial, isto é, um texto central cujas expansões são hierarquizadas, ou pode ter o formato de rede, ou seja, segmentos de informação interligados por possibilidades de acesso e relações múltiplas. Tanto a forma de organização textual axial como a em rede fraturam a sequencialidade do texto tradicional, exigindo, portanto, novas estratégias de leitura e uma nova relação do leitor com o texto, uma vez que ele terá que fazer escolhas e determinar a ordem de acesso aos diferentes segmentos disponibilizados no hipertexto e o eixo coesivo que confere um sentido global ao texto lido (BRAGA, 2007).

Não se pode deixar de levar em conta também que, uma vez que praticamente todos os gêneros textuais podem ser lidos em camadas na tela do computador, é criada uma continuidade textual que se abre em cascata para o leitor. Esse aspecto pode provocar uma inquietação no leitor por não ter à mão a materialidade do papel ou livro, com a qual está acostumado, além de ele ter que enfrentar outros critérios para "distinguir, classificar e hierarquizar discursos" (CHARTIER, 2002, p. 23).

Não que o tipo de organização textual na qual há uma ruptura da sequencialidade seja exclusiva dos textos digitais (as divisões textuais em seções, notas de rodapé e experiências literárias já fazem uso de tal recurso), mas, para Landow (1997), a multiplicidade do hipertexto é uma chamada para um leitor mais ativo e atento ao percurso que ele próprio escolheu trilhar. Marcuschi (2006) entende o hipertexto como um atualizador de recursos que antes já eram usados nos textos escritos, mas que com as tecnologias de comunicação e informação foram ressignificados, surgindo uma nova forma de textualidade. O linguista nos relembra que o conceito de hipertexto sempre existiu na tradição ocidental, mas que com o hipertexto foi possível relacionar notas, citações, bibliografias, referências, imagens e sons de forma mais

integrada e, portanto, eficiente. Em suma, o hipertexto "subverte os movimentos e redefine as funções dos constituintes textuais clássicos" (MARCUSCHI, 2006, p. 1).

Uma grande contribuição de Landow (1997) para a significação do hipertexto diz respeito ao estrangulamento da tirania univocal. Segundo o estudioso,

Como os leitores se movem por uma teia ou rede de texto, eles continuamente mudam o centro e, portanto, o foco ou princípio organizador de sua investigação e experiência. O hipertexto, em outras palavras, proporciona um sistema infinitamente recentrável, cujo ponto provisório de foco depende do leitor, que se torna um leitor verdadeiramente ativo em outro sentido. (LANDOW, 1997, p. 36).<sup>2</sup>

Uma das características fundamentais do hipertexto é ele ser composto por textos que não identificam um eixo principal na sua organização. Esse é o princípio de mobilidade dos centros (COELHO; ALEXANDRIA, 2010, p. 132 apud LÉVY; AUTHIER, 2000), que prevê a simultaneidade de diferentes centros de organização da leitura, que se modificam e se reorganizam a cada clique.

O hipertexto deixa de ter um princípio de organização principal, o que permite ao leitor estabelecer focos provisórios no texto, tornando-o ativamente participante da atividade de leitura. Nesse novo princípio de organização textual, o leitor não fica preso a uma hierarquia imposta pelo autor do texto, embora, obviamente, este, na produção textual, dê direcionamentos de leitura. Entretanto, é o leitor que decide se irá ou não segui-los. Sobre esse ponto, Ferreira e Fellipe (2010) acreditam que

<sup>&</sup>lt;sup>2</sup> Tradução minha do original em inglês: "As readers move through a web or network of text, they continually shift the center-and hence the focus or organizing principle-of their investigation and experience. Hypertext, in other words, provides an infinitely recenterable system whose provisional point of focus depends upon the reader, who becomes a truly active reader in yet another sense".

ao leitor cabe um papel semelhante ao do autor, pois àquele é permitido organizar trechos de leitura enquanto clica nos *links* sem levar em conta a proposta e a intenção do autor.

Xavier (2010), ao discutir essa relação autor e leitor de (hiper)textos, diz que o autor deixa na tela sugestões de caminhos a serem percorridos, os quais são totalmente violáveis, pois é princípio norteador na construção de hipertextos a otimização das escolhas de trilhas no espaço digital, a qual vai garantir ao leitor um leque amplo de opções de checagem. Entretanto, o linguista chama atenção para o fato de que a liberdade de escolha do leitor é condicionada às decisões do autor do hipertexto sobre a disponilização dos links, sendo estes, portanto, um meio de respaldar o ponto de vista do seu produtor. Essa perspectiva não é naturalmente nova, visto que as pistas textuais e as pistas de contextualização (GUMPERZ, 1998 e 1999; MENEGUELLI, 2012), tanto na produção escrita quanto na produção oral, já foram investigadas como pontos de passagem do autor.

Mesmo havendo algumas nuances diferenciais entre pontos de vista, o que há de convergente é a reconfiguração do papel do leitor em ambiente digital. A leitura em si sempre foi um ato corporal: mexer os olhos, imaginar uma cena, folhear as páginas com os dedos, segurar o livro. Diante da tela do computador, a leitura continua sendo ativa, mas provocando outros sentidos. O olhar e os ouvidos são direcionados para lugares diversos: cartazes publicitários, hiperlinks que remetem a temas não relacionados a uma leitura prévia, o som de chamada para uma conversa em um chat, entre outras ações que podem ser exigidas do leitor, que, cognitivamente, precisa saber administrar as alternativas que a tela do computador lhe apresenta e que impõe a ele tomar uma decisão.

# 3 - Artefatos semióticos na composição de sentidos

Para Lemke (2002), a representação visual e a linguagem escrita têm coenvolvimento cultural e histórico que complementam e completam

uma a outra de forma coordenada e integrada. Nas práticas de construção humana de significação, elas são inseparáveis na maioria das vezes. Tentando entender as possibilidades semânticas de textos hipermodais, Lemke lança os seguintes questionamentos: Que significados podem ser construídos através da combinação de signos verbais, visuais e de outros sistemas semióticos? Como podemos interpretar os significados de componentes em complexos multimodais? Quando combinamos texto e imagem, cada imagemtexto passa a ser o produto multiplicativo do conjunto de textos linguísticos possíveis e do conjunto de imagens possíveis. Ora, nenhum texto verbal corresponde plenamente a uma imagem. Nenhuma imagem ou representação visual representa efetivamente um texto verbal. E é essa incomensurabilidade essencial que permite que novos significados sejam criados de combinações entre as modalidades com cada linguagem salientando um conteúdo informacional conforme o seu potencial comunicativo.

Na composição de um texto em meio digital, a coesão hipertextual é tecida por artefatos semióticos ligados por redes complexas (os links) que geram significados em diferentes escalas de organização sintagmática. Vejamos, a seguir, na 10ª semana da disciplina Geografia II do curso de Pedagogia na modalidade à distância, como um problema na coesão do texto pode afetar a leitura do estudante:

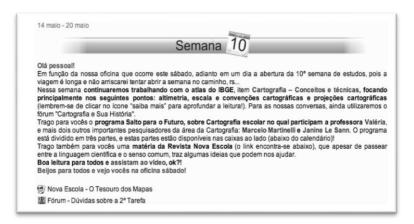


Figura 1: Coesão textual

Na semana 10 (Figura 1), verificamos um problema de referenciação em "(lembrem-se de clicar no ícone 'saiba mais' para aprofundar a leitura!)", que poderia ser solucionado com a disponibilização de um hyperlink referenciado verbalmente. Não há um hyperlink na semana 10 que expanda a informação que deveria estar contida no ícone "saiba mais", que também não foi localizado nas semanas anterior e nem posterior. Dessa maneira, foi criada uma expectativa nos estudantes que não pôde ser constatada devido a um problema de coesão interna que impediu a ligação entre partes do texto.

A pluralidade dos contextos atravessados pelas mensagens hipermodais exige o conhecimento aberto de novas estratégias de coesão e coerência textuais, uma vez que a referência dos *links* deve ser disponibilizada corretamente, a fim de que o leitor não se perca no ciberespaço, e a ampliação da informação deve ser visada partindo da consideração do autor acerca da coerência interna que os links devem estabelecer.

Outros elementos que estariam para Lemke associados a estratégias de produção escrita em hipertextos são o uso de cor e a organização estrutural de unidades textuais em cabeçalhos, títulos, links - como, por

exemplo, em um site, abaixo do cabeçalho, habitualmente há uma página coesa com título ou legenda. Os significados *orientacionais* são dados através da combinação entre visual e verbal, a fim de produzir uma orientação para o leitor sobre o conteúdo temático apresentado.

Lemke (2002a) diz que a hipermodalidade é mais do que a multimodalidade, porque ela é a confluência da multimodalidade e da hipertextualidade. A hipermodalidade não é simplesmente uma justaposição do verbal, do visual e do som; nela são projetadas interconexões múltiplas entre os textos e suas semioses, tanto potencial quanto explicitamente. Não só há as ligações entre as unidades de textos (lexias), mas também há ligações entre elementos verbais, visuais e sonoros.

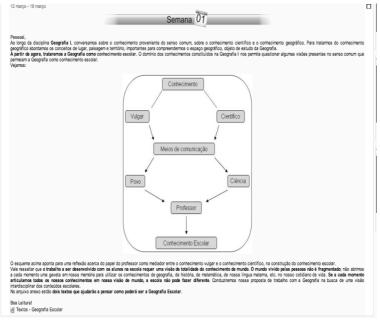


Figura 2: Composição hipermodal

A professora/tutora da disciplina usa os recursos visuais de cores e negrito para dar destaque aos tópicos que julga serem relevantes para o aluno, tais como: conhecimento escolar, visão de totalidade do conhecimento do mundo e que o mundo vivido não é fragmentado. De acordo com Lemke (2002a), como já vimos, a representação visual e a representação verbal complementam e completam uma a outra se usadas de forma coordenada e integrada na construção de significados. Para o semiólogo, a combinação de sistemas semióticos produz novos tipos de significados mais específicos e que não estariam disponíveis isoladamente. Também Braga (2004), nesse sentido, diz que, como os significados veiculados pelas diferentes modalidades de linguagem não são independentes entre si, as possibilidades de construção de sentido se ampliam em relação ao uso de uma única modalidade, uma vez que cada uma delas se complementa, auxiliando na interpretação geral.

Verificamos na Figura 2 que essa integração não só ocorre, como também ressignifica a informação quando é dado um destaque verbovisual à posição inicial defendida de que a Geografia será tratada na disciplina como um conhecimento escolar, como pode ser visto na relação de destaque em: a partir de agora trataremos a Geografia como conhecimento escolar. O recurso visual tem, nesse caso, a função de reforçar a posição da professora/tutora. Para destacar que os conteúdos da disciplina Geografia I e II mantêm um elo de continuidade temática, a professora/tutora apresenta um esquema visual sobre como o conhecimento circula no mundo até chegar ao professor, mediador na escola do conhecimento de senso comum e do conhecimento científico e responsável por articulá-los e transformá-los em conhecimento escolar.

No desenho da aula (Figura 2), o tema da aula "<u>a Geografia como</u> <u>conhecimento escolar</u>" remete o leitor ao que está mais evidente com os recursos visuais, os quais deixam em saliência que o professor é o mediador do conhecimento escolar. A expressão verbal, cujo destaque em vermelho desvia o olhar do leitor do esquema visual para ela, enfatiza a importância do conhecimento escolar para o tema da aula.

Para compreender o tema e participar da aula, o aluno deve fazer uma leitura baseada no *interactive model*, uma vez que ele precisa fazer relações entre as informações visuais e verbais, além de inferir o significado do esquema. Caso o aluno não realize esses percursos cognitivamente, sua compreensão da aula pode ser prejudicada.

É dado um destaque maior ao esquema visual, no nível organizacional, por estar ele localizado no centro da tela, visando direcionar o olhar do leitor primeiramente para a informação visual. O esquema visual também revela a estratégia da professora/tutora em dar destaque a que tipo de conhecimento o professor deve mediar, reforçando visualmente o texto verbal.

As escolhas da professora/tutora pelo uso combinado de informações verbo-visuais e a forma de disponibilizá-las na aula da semana 1 (Figura 2) produz no aluno uma significação de importância das informações apresentadas visualmente, tanto no esquema como nos destaques dados pela cor. Essa integração verbo-visual organiza as informações relevantes para o estudante, orientando a sua leitura para o conteúdo de destaque, ao qual ele deve se ater.

O texto que segue o esquema visual é uma posição que se repete em outro signo, o verbal, reafirmando a posição inicial de que ao professor cabe articular os conhecimentos científico e vulgar para transformá-los em conhecimento escolar.

Para dar mais um reforço a sua tese, a tutora apresenta uma posição associada disponibilizada no *hiperlink* <u>Textos – Geografia Escolar</u>, no qual o aluno poderá aprofundar a sua leitura a partir de um texto produzido por um especialista da área, sendo esse tipo de evidência empírica um argumento de autoridade.

A aplicação do letramento digital é um caminho ético para a inserção de sujeitos no mundo social, visto que uma parte considerável das práticas linguageiras, contemporaneamente, acontece no ciberespaço. O

letramento digital deve almejar os letramentos críticos pela análise de valores, intenções, estratégias e efeitos de sentido provocados pelo autor de textos nos quais foram usados recursos hipermodais. O hipertexto anuncia de forma indicial os recortes sinalizados pelo autor, que também pode salientar em mais de um sistema semiótico aspectos que julga relevantes.

Sobre esse aspecto da saliência que cada modo pode sugerir, Kress (1996) mostra a relação entre verbal e imagem, argumentando contrariamente que o verbal é o modo responsável por carregar todo o conteúdo informacional e a imagem o repete.

As estruturas visuais produzem significado, assim como as estruturas linguísticas. Entretanto, as interpretações das experiências são diferentes, além de gerarem formas de interação social diferentes. Os significados que podem ser compreendidos pela linguagem e pela comunicação visual coincidem em parte, isto é, algumas coisas podem ser expressas tanto visualmente quanto verbalmente; e em parte elas divergem – algumas coisas podem ser "ditas" somente visualmente, outras somente verbalmente. Mas, mesmo quando algo pode ser "dito" tanto visualmente quanto verbalmente, a forma como será expressa é diferente (tradução minha).

Estaria havendo um certo nível de especialização das linguagens (KRESS e VAN LEWUVEN, 1996, p. 3) e essa especialização, esse "como" os modos significam, devem ser motivo de atenção em uma prática de letramento.

O letramento digital, como conhecimento de práticas sociais mediadas pelo uso do computador, pode ativar no leitor o *interactive model*, pois ele precisa acionar as suas experiências prévias de leitura oriundas tanto de textos digitais quanto de textos impressos e, também, manter-se centrado no texto, já que lhe são demandadas ações para prosseguir a

leitura e capacidade inferencial e relacional, porque precisa gerenciar várias camadas de informações.

#### 4 - Conclusões

Procurei apresentar teorias de leitura e produção de textos com o objetivo de incentivar a discussão da importância de serem adotadas no âmbito escolar práticas de letramento digital que ativem o interactive model. Como já dito, este texto tem como objetivo principal analisar estratégias de leitura e de escrita de textos digitais, uma vez que os hipertextos e os textos multissemióticos têm circulado nas mais diversas esferas sociais. Por isso, a escola e outras instituições devem se preparar para adotar o letramento digital.

Principalmente à escola cabe o papel de repensar práticas pedagógicas de leitura e de escrita que priorizem o aspecto multimodal dos textos, abrangendo análises sobre as especificidades das várias linguagens utilizadas na produção textual, a integração entre elas na geração de múltiplas possibilidades de (re)construir significados e o hipertexto como tecnologia enunciativa (XAVIER, 2002). Essa prática não pode deixar de levar em conta a instância de produção de (hiper)textos, pois o sujeito-autor, ao integrar recursos multimodais no hipertexto, deixa marcas ideológicas na sua composição textual. Embora o leitor seja o responsável por reconstruir as trajetórias iniciadas pelo autor, é este quem dá início ao percurso, moldando-o a partir das suas ideologias. Conforme propõe Lemke, a comunicação multimodal é um recurso poderoso porque pode criar significados que vão permitir às pessoas ler e escrever novas formas de expressão de maneira mais crítica.

### **REFERÊNCIAS**

BARTON, D. & HAMILTON, M. *Local Literacies*: Reading and writing in one community. London, Routledge,1998.

BRAGA, D. B. Letramento na internet: o que mudou e como tais mudanças podem afetar a linguagem, o ensino e o acesso social. In: Ângela Kleiman,

Marilda C. Cavalcanti. (Orgs.) Lingüística aplicada: suas faces e intrefaces.
Campinas: Mercado das Letras, 2007.
A comunicação interativa em ambiente hipermídia: as vantagens
da hipermodalidade para o aprendizado no meio digital. In: L. A. Marchuschi e
A. C. Xavier. (Orgs.) <i>Hipertexto e gêneros digitais</i> . Rio de Janeiro: Lucerna, 2004.
CAVALCANTE, M. C. B. Mapeamento e produção de sentido: os links no
hipertexto. Hipertexto e gêneros digitais: novas formas de construção do
sentido. São Paulo: Cortez, 2010.
CHARTIER, Roger. <i>Os desafios da escrita</i> . São Paulo: Editora Unesp, 2002.
COELHO, R.; ALEXANDRIA, M. Para uma retórica do hipertexto. Revista
<i>Ipotesi</i> , Juiz de Fora, v. 14, n. 1, p. 127-141, jan./jun. 2010.
COSCARELLI, C. V.; RIBEIRO, A. E. (Orgs.) Letramento Digital: aspectos sociais
e possibilidades pedagógicas. Belo Horizonte: Ceale, Autêntica, 2007.
DONDIS, A. D. Sintaxe da linguagem Visual. Trad. Jefferson Luiz Camargo. São
Paulo: Martins Fontes, 2007.
GRAFF, H. J. <i>Os Labirintos da alfabetização</i> . Porto Alegre: Artes Médicas, 1995.
GUMPERZ, John. Convenções de contextualização. In: RIBEIRO, Branca Telles;
GARCEZ, Pedro M. (Orgs.) Sociolingüística interacional: antropologia,
lingüística e sociologia em análise do discurso. Porto Alegre: AGE, 1998.
GUMPERZ, John. On Interactional sociolinguistic method. In: SARANGI,
Srikant; ROBERTS, Celia (Eds). Talk, work an institucional order. Berlin, New
York: Mouton de Gruyter, 1999.
KATO, M. A. <i>O aprendizado da leitura</i> . 4ª ed. São Paulo: Martins Fontes, 1995.
KLEIMAN, Ângela. Os significados do letramento. Campinas, São Paulo:
Mercado de
Letras,1995.
KRESS, G. Visual and verbal modes of representation in electronically
mediated communication: The potentials of new forms of text. In: SNYDER, I.
(ed.). Page to screen: Taking literacy into the electronic era. London, New York:
Routledge, 1998, p. 53-79.
Literacy in the new media age. London, New York: Routledge, 2003.
Gains and Losses. In: Computers and composition. Volume
22/No.1/2005. pp.5-23, 2005.
,VAN LEEUWEN, T. Reading images: the Grammar of Visual Design.
London: Routledge, 1996.
; Front Pages The (Critical) Analysis of Newspapper Layout. In:
BELL, A. and GARRET, P. (Eds.) Approaches to media discourse, pp. 186-219.
London: Blackwell, 1998.

Acesso em: 10 out. 2010.

LANDOW, George P. Hypertext 2.o. Hypertext: the convergence of contemporary Critical Theory and technology. Maryland: The John Hopkins University Press, 1997. LEMKE, J.L. *Travels in hypermodality*: visual communication 1(3), 2002a. \_\_\_\_\_. LEMKE, J. Important theories for research topics on this website. [2000?]. Disponível em: http://academic.brooklyn.cuny.edu/education/jlemke/ theories.htm#General%20Semiotics. Acesso em: 24 mar. 2010. \_\_. Multimedia Genres for Scientific Education and Science Literacy. In M. J. Schleppegrell & C. Colombi, Eds., Developing advanced literacy in first and pp.21-44. 2002b. second languages. Erlbaum. Disponível http://www.jaylemke.com/storage/MultimediaGenres-Science-2002.pdf. Acesso em: 14 mar. 2011. \_\_\_\_\_. Multiplying Meaning: Visual and Verbal Semiotics in Scientific Text. In: J.R. Martin & R. Veel, Eds., Reading science. London: Routledge. (pp.87-113). 1998. LÉVY, Pierre. Cibercultura. São Paulo: Ed. 34, 1999. MARCUSCHI, L. A. Gêneros textuais emergentes no contexto da tecnologia digital. In: \_\_\_\_\_\_; XAVIER, A. C. S. (Orgs.). Hipertexto e gêneros digitais: novas formas de construção do sentido. São Paulo: Cortez, 2010, p. 15-80. \_\_\_\_\_. O hipertexto como um novo espaço de escrita em sala de aula. In: Linguagem & ensino, Vol. 4, No. 1, 2001, p. 79-111. Disponível em: http://www.pucsp.br/~fontes/ln2sem2006/f\_marcuschi.pdf. Acesso em 29 mai. 2011. Linearização, cognição e referência: o desafio do hipertexto (2006). http://www.pucsp.br/~fontes/ln2sem2006/17Marcus.pdf. Disponível

MENEGUELLI, Gisella. Pistas de Contextualização em Negociações de Contexto: Uma Análise Linguística no Contexto Interacional Fórum de Discussão. In: Petrov, P.; Sousa, Pedro Q. de; Samartim, Roberto López-Iglésias; Torres Feijó, Elias J. (eds.): *Avanços em ciências da linguagem*. Santiago de Compostela - Faro, Associação Internacional de Lusitanistas: Através Editora, 2012, pp. 331-342.

OLIVEIRA, S. Texto visual e leitura crítica: o dito, o omitido, o sugerido. *Linguagem e ensino*. Pelotas, vol. 9, No.1, jan/jun. 2006, p.15-39.

ROJO, Roxane. Letramentos múltiplos, escola e inclusão social. São Paulo: Parábola Editorial, 2009.

SANTAELLA, Lúcia. *Navegar no ciberespaço*: o perfil do leitor imersivo. São Paulo: Paullus, 2004.

SNYDER, I. (ed.). *Pagetoscreen*: Taking literacy into the electronic era. London, New York:Routledge, 1996.

SOARES, Magda. *Letramento*: um tema em três gêneros. Belo Horizonte, Minas Gerais: Autêntica, 1998.

XAVIER, Antonio C. S. O Hipertexto na sociedade da informação: a constituição do modo de enunciação digital. Tese de Doutorado, Unicamp: inédito, 2002.

——. Letramento digital e ensino. s/d. Disponível em http://www.ufpe.br/nehte/artigos/Letramento%2odigital%2oe%2oensino.pdf. Acesso em 25 out. 2010.

# PARTE II FERRAMENTAS E REPOSITÓRIOS DIGITAIS

# O uso de elementos semânticos no processo de recuperação da informação em ambientes digitais

José Eduardo Santarém Segundo<sup>1</sup>

# Apresentação

O capítulo que apresento neste livro é fruto da palestra, com mesmo título, apresentada no II Simpósio Internacional e VI Simpósio Nacional de Literatura e Informática, organizado pelo NuPILL - Núcleo de Pesquisas em Informática, Literatura e Linguística, na Universidade Federal de Santa Catarina.

A estrutura deste capítulo resume parte da pesquisa, em andamento, deste pesquisador que estuda a viabilização de técnicas e elementos semânticos na caracterização de uma estrutura semântica para repositórios digitais.

A pesquisa iniciou-se com o desenvolvimento de um modelo teórico que resultou na tese de doutorado "Representação Iterativa: um modelo para repositórios digitais". A tese foi premiada como melhor tese na área pela ANCIB (Associação Nacional de Pesquisa e Pós-Graduação em Ciência da Informação), no ano de 2011. Recebeu ainda a menção honrosa no Prêmio Capes de Tese na área de Ciências Sociais Aplicadas, também em 2011.

\_

<sup>&</sup>lt;sup>1</sup> Professor e coordenador do Curso de Graduação em Ciências da Informação e da Documentação, do da Universidade de São Paulo (USP) em Ribeirão Preto e professor colaborador no Programa de Pós-Graduação em Ciência da Informação da Universidade Estadual de São Paulo (PPGCI/UNESP) em Marília. CONTATO: santarem@usp.br

O modelo teórico vem ganhando traços de implementação e novos testes para que seja homologado por completo em breve, porém ainda não há previsão de quando a pesquisa seja realmente finalizada com todos os itens descritos no modelo teórico.

A motivação originária para o desenvolvimento desta pesquisa nasceu da afinidade deste pesquisador com os repositórios digitais e principalmente pela forte relação com o uso da plataforma *Dspace*. Também foi motivador o contato com a Web Semântica, que vem sendo pesquisada e abordada desde o ano de 2004, quando ainda finalizava o mestrado usando este tema.

Portanto a proposta apresentada neste capítulo tem como objetivo principal abordar o contexto dos ambientes digitais, com ênfase para os repositórios digitais, de forma que se busque evidenciar o uso de elementos semânticos como facilitador no processo de armazenamento e recuperação da informação, através do processo de Folksonomia Assistida e do modelo Representação Iterativa.

# Introdução

Grande parte do conteúdo textual e imagético de todo o mundo tem sido construído em formato digital. Com raras exceções, como as cadernetas que marcam o "fiado" da quitanda, receitas de bolo, algumas prescrições e prontuários médicos, contratos e um ou outro tipo de documento que ainda persistem no formato analógico. Armazenar documentos em formato não eletrônico deixou de ser uma tarefa trivial no cotidiano das pessoas.

Em um passado recente ainda era possível notar que grande parte dos documentos já eram produzidos de forma digital, mas ainda continuavam tramitando de forma impressa. Não era raro observar as pessoas carregando documentos que haviam sido elaborados em

formato digital, mas que eram transformados em impresso para que pudessem chegar a seus destinatários.

Não há como negar a presença das tecnologias de informação e comunicação (TIC's) nos modelos de gerenciamento de informação atuais. A Internet tornou-se o principal meio condutor de informações e a informação digital tem se tornado muito mais rápida e efetiva, em comparação com a informação em suporte impresso.

A explosão do uso da Internet, do armazenamento em nuvem e principalmente o uso de dispositivos móveis, tem mudado de forma drástica a maneira de construir, armazenar e principalmente recuperar conteúdo informacional. Essa mudança nos processos que envolvem o fluxo informacional são muito representativas e nos apresentam um modelo totalmente diferente de lidar com a informação. Ela deixou de ser analógica e passou para o modelo digital.

Atualmente é possível verificar que instituições e organizações variadas têm tomado decisões a respeito de processos que envolvem o uso da informação em formato digital. Nota-se, inclusive, que grande parte dessas instituições tem se preocupado em fazer com que os documentos que ainda persistem em continuar no formato papel voltem a ser convertidos para o formato digital através de processos de digitalização.

No que tange à produção de documentos em formato papel, ele se torna inviável em vários sentidos, desde a produção, quando exige alto custo de papel e toner/tinta para imprimir, caracterizando-se contra os atuais modelos de sustentabilidade, passando pela guarda que exige cada vez mais espaço físico e indo até o processo de recuperação que em muitos casos envolve perda de tempo e principalmente impossibilita que o conteúdo de um documento seja compartilhado entre várias pessoas ao mesmo tempo (SANTAREM SEGUNDO; DONADI, 2013).

Esse cenário indica claramente que a Ciência da Informação, responsável por pesquisas que envolvem estruturas de dados e fluxos

informacionais, tem um papel preponderante na construção de modelos que facilitem o acesso à informação nesse novo formato digital.

É por todo esse novo contexto, essa mudança de paradigma, que a organização e a recuperação da informação, principalmente quando esta se encontra na Internet, tem sido pauta de muitas das pesquisas que envolvem tecnologia da informação. Entretanto, esses desafios não param por aí, a partir do momento em que se produz uma quantidade exacerbada de informação em formato digital, é necessário que essas informações estejam não apenas disponíveis, mas também integradas com documentos e objetos digitais que têm relação direta com o conteúdo, através de relações semânticas.

Considera-se relação semântica entre dois dados distintos qualquer tipo de ligação realizada pelo cérebro humano que possa associar estes dados a partir de um determinado cenário.

Ressalta-se que, neste texto, o uso de elementos e processos que tratam da questão semântica serão sempre aplicados em ambientes digitais do tipo repositório.

A abordagem pelos repositórios é dada devido à inserção que esta ferramenta tem tido no contexto atual. As instituições têm optado por modelos de gestão de conteúdo digital que primam pelo uso de técnicas difundidas na Ciência da Informação, como a descrição, o uso de metadados para interoperabilidade, a indexação, o uso de vocabulários controlados, entre outros, e para isso se apoiam no uso de repositórios.

# Repositórios digitais

Repositórios digitais são ambientes de informação constituídos sobre ferramentas/plataformas tecnológicas, que também podem ser chamadas de sistema de informação, capazes de receber depósitos de objetos digitais, em vários formatos, sejam eles realizados pelos próprios autores ou por equipes treinadas para este fim, de modo a

armazenar e organizar estes objetos para que possam ser recuperados e principalmente preservados por longa data.

Os repositórios digitais, sejam eles temáticos ou institucionais, apresentam características semelhantes, possuem uma estrutura comum de submissão e acesso às informações e são desenvolvidos segundo padrões de interoperabilidade específicos, que potencializam o uso desses sistemas para agregação e divulgação da informação digital. Algumas das características ou observações quanto à estrutura destes repositórios são impostas pelos padrões que utilizam (SANTARÉM SEGUNDO, 2010, tese).

De acordo com os objetivos propostos, um repositório digital pode contemplar uma infinidade de tipos de documentos, ou seja, ter uma tipologia variada de documentos. Mesmo com a concepção de que os repositórios foram desenvolvidos para divulgar documentos já publicados nos meios tradicionais, não há uma delimitação sobre os tipos de documentos que atualmente podem compor um repositório (CAFÉ, 2003).

Bekaert e Van de Sompel (2006) definem repositórios como:

conjuntos de documentos coletados, organizados e disponibilizados eletronicamente. No contexto específico dos repositórios, os documentos adquirem novas configurações e são denominados objetos digitais ou estrutura de dados digitalmente codificados, composta pelo conteúdo de informação, metadados e identificador.

Difundidos principalmente na organização de material científico e utilizados em larga escala pelas universidades e institutos de pesquisa, conforme vimos nas definições apresentadas, os repositórios têm chamado também atenção de outros tipos de organização, devido à capacidade de gerenciar grandes volumes de dados facilitando o processo de armazenamento e recuperação da informação.

Para a implementação técnica são vários os softwares/plataformas disponíveis, tanto em iniciativas de software livre, open-source e até alguns que sugerem a aquisição de uma licença de uso. Entre os principais, atualmente encontram-se: Dspace , ,GNU E-prints , OPUS , Open Repository , DiVA , Fedora , Omeka .

Durante esta pesquisa ficará evidente a necessidade de agregar elementos e recursos de Web 2.0 e de Web 3.0 nas ferramentas que implementam repositórios, visto que não é possível identificar esses elementos nas versões originais disponibilizadas.

Da necessidade de uso dos recursos de Web 3.0, principalmente os elementos, linguagens e tecnologias apresentadas como essenciais no "bolo de noiva" da Web Semântica, desenvolvido por Tim Berners-Lee, e da folksonomia, elemento básico da Web 2.0 é que se constrói grande parte desta pesquisa.

#### **Folksonomia**

Folksonomia é a tradução do termo criado por Thomas Vander Wal, a partir da junção das palavras *folk* (povo) com *taxonomy* (Taxonomia). Wal (2006, p. 1) define folksonomia como "resultado de atribuição livre e pessoal de *tags* (etiquetas) a informações ou objetos (recursos na *web*), visando a sua recuperação".

A folksonomia é considerada elemento fundamental na construção do modelo Representação Iterativa.

O propósito principal da folksonomia neste contexto é permitir que usuários comuns criem *labels/tags* que possam descrever ou apontar para o conteúdo que estão inserindo durante o autoarquivamento de documentos digitais em repositórios digitais, de modo que os recursos possam ser recuperados posteriormente pelo próprio usuário ou ainda por outros usuários que procurem informações no ambiente digital.

Entende-se que a folksonomia se caracteriza como uma forma de inserir e relacionar recursos por meio da descrição dos mesmos através de palavras-chave, de forma aberta, que tem como principal objetivo facilitar o processo de gerenciamento e recuperação das informações em ambientes digitais.

O fato da folksonomia promover a participação do usuário de forma livre, permite que a criação das *tags* receba o nome de vocabulário descontrolado, em uma alusão aos vocabulários controlados, que são um recurso disponível para alinhar indexação de informação dentro de um conjunto de palavras fixas que representam um determinado domínio de informação.

Aquino (2007, p. 10) faz essa abordagem:

Poderíamos dizer que a folksonomia é um tipo de vocabulário descontrolado. Isso não quer dizer que o esquema seja uma desordem total [...] Na verdade, trata-se de um mecanismo de representação, organização e recuperação de informações que não é feito por especialistas anônimos, o que muitas vezes pode limitar a busca por não trazer determinadas palavras-chave, mas sim um modo onde os próprios indivíduos que buscam informação na rede ficam livres para representá-la, organizá-la e recuperá-la, realizando estas ações com base no senso comum.

A folksonomia mudou o paradigma em relação à recuperação da informação em ambientes *Web*, tanto que é comum ver sites apresentando buscas baseadas em palavras-chave que foram inseridas pelo próprio usuário dentro do ambiente. Portanto, trata-se de um recurso rico da Web 2.0, que contribui de forma acentuada para o fortalecimento e solidificação da Internet como plataforma para construção de informação coletiva.

# Representação Iterativa

A Representação Iterativa caracteriza-se como um conjunto de técnicas formando um modelo de estrutura para adaptação dos repositórios digitais, para que estes possam efetivar o relacionamento de termos e a recuperação semântica da informação.

Dentro deste contexto, a Representação Iterativa deve transformar um repositório digital científico em uma ferramenta apta a descrever, armazenar e recuperar informação, permitindo a recuperação semântica e a construção coletiva de uma estrutura relacional semântica de informações por meio de Folksonomia Assistida; e as técnicas utilizadas no desenvolvimento da estrutura sugerida (SANTAREM SEGUNDO, 2010).

A atividade de inserir informações em um repositório digital compreende o processo de inicialmente descrever o conjunto de informações que representa os metadados do objeto a ser inserido e, na sequência, realizar o envio do arquivo principal e também dos arquivos complementares, se houverem, para que todo o conjunto de informações seja armazenado no repositório.

Como já fora mencionado, a pesquisa sugere a construção de um modelo estrutural para repositórios digitais, de forma que esses ambientes possam agregar funcionalidades que atuem no sentido de garantir ao usuário uma melhor interface de comunicação com o sistema e ainda evoluir no processo de recuperação da informação, possibilitando a apresentação de resultados baseados em relação semântica, em associação de conteúdos, e não apenas em comparação sintática, como é realizado atualmente.

O modelo — Representação Iterativa — parte do princípio de que o usuário deverá ter uma interface diferente para inserção de dados no repositório digital. A princípio, a única alteração em relação à interface padrão de descrição do recurso será no momento de informar as

palavras-chave, visto que estes campos deverão vir com uma informação de que, além de configurar como palavras-chave, os dados descritos ali serão também utilizados como *tags* (SANTAREM SEGUNDO, 2010).

O fato de caracterizar o uso de *tags* já cria neste ambiente a ideia de que o repositório tratará as palavras-chave como parte da concepção de folksonomia, e, portanto, deverá implementar recursos que permitam a recuperação da informação em novos formatos, como uma nuvem de *tags*, por exemplo.

No momento em que o usuário iniciar o processo de descrição da *tag* deverá ocorrer uma intervenção do sistema, de forma que se caracterize um processo que denominamos Folksonomia Assistida.

Folksonomia Assistida é um processo de apoio ao usuário, no momento de definir os termos mais adequados para as *tags* que referenciarão seu trabalho depositado em um repositório digital. O processo é composto por duas partes principais (SANTAREM SEGUNDO, 2010).

A primeira parte implica que, para a implementação da Folksonomia Assistida, deverá ser alterada a interface de comunicação do usuário com o repositório, ou então desenvolvida uma nova interface, para a inserção de informações no campo palavra-chave, utilizado como referência para a inserção de conteúdo para as *tags*.

Nesse primeiro passo, deve-se apresentar ao usuário, no momento da digitação da *tag*, de forma sistemática, um conjunto de informações já previamente inseridas no sistema, como uma sugestão de *tags*. A busca de informação para fazer a sugestão é baseada em busca sintática.

O recurso de "sugestão" foi inicialmente apresentado pelo Google em sua ferramenta de busca, porém neste contexto tem sido adaptado para facilitar o processo de descrição do recurso pelo usuário.

O conjunto de informações que deverá ser apresentado ao usuário no momento que este estiver digitando será baseado nas *tags* já inseridas no sistema e também nos termos que fazem parte de uma estrutura de representação do conhecimento das áreas de especialidades, que deverá estar associada ao repositório como parte do modelo estrutural proposto. Esta estrutura deverá ser caracterizada como uma ontologia.

Assim que o usuário descrever as *tags*, aceitando ou não as sugestões, o sistema receberá a informação e dará início a um segundo passo para a concepção da Folksonomia Assistida.

No segundo passo, o repositório deverá receber os termos enumerados pelo usuário e proceder à pesquisa de relacionamento da informação dada pelo usuário em relação ao conjunto de informações internas que a ferramenta dispõe.

O processo de relacionamento em questão é justamente uma busca de relações dentro de uma ontologia, que são instrumentos que permitem uma busca hierárquica de termos em formato vertical e horizontal. Nesta pesquisa utilizar-se-á uma ontologia, descrita em linguagem OWL, principalmente por ser uma linguagem que vem sendo aprimorada constantemente, e conta com indicação de uso pelo W<sub>3</sub>C.

A busca por termos relacionados em uma ontologia escrita com a linguagem OWL deve ser realizada com a linguagem Sparql, que tem como princípio justamente recuperar informações relacionadas em uma linguagem para descrição de ontologias.

Esse segundo passo da Folksonomia Assistida, além de recuperar termos relacionados de forma semântica em uma estrutura de representação do conhecimento das áreas de especialidades, deverá também buscar informações no conjunto de *tags* já inseridas no sistema, principalmente em seus relacionamentos horizontais. A busca por termos na estrutura de representação do conhecimento deverá acontecer em níveis pré-estabelecidos pelo administrador do ambiente,

e a busca por relacionamentos horizontais no conjunto de *tags* já descritas também poderá ser mediada pelo administrador, que deverá informar a quantidade de termos oferecidos para cada termo digitado pelo usuário. Esses conceitos poderão ser previamente parametrizados e adaptados conforme o repositório for sendo ampliado com novos depósitos.

A seguir, após essa busca interna por relacionamentos em relação ao termo descrito pelo usuário, o sistema apresentará novamente ao usuário um conjunto de termos que poderão ser aceitos de forma total ou parcial, ou ainda descartados pelo usuário, como sugestão final de tags para o recurso a ser inserido. Em todo esse processo, cabe ao usuário decidir as tags que melhor representem seu recurso digital dentro do domínio do repositório digital científico em que está sendo realizado o depósito.

A utilização de termos de uma estrutura de representação do conhecimento e também de *tags* já inseridas no sistema não tem o objetivo de engessar a criatividade do usuário, nem tampouco de descaracterizar o termo folksonomia, pois o sistema permite claramente que o usuário decida livremente os termos que deverão ser utilizados como *tags*. A Folksonomia Assistida tem como principal característica oferecer ao usuário um conjunto de termos que já estão sendo empregados no sistema, de forma que ele possa usar a base de conhecimento do próprio repositório para qualificar a descrição de seu recurso.

Santarém Segundo (2010, p. 276) afirma que "Folksonomia Assistida prima pela consistência das *tags*, de forma que o usuário do sistema evite abreviações, plurais/singulares ou ainda palavras que possam dificultar a recuperação da informação, posteriormente".

O processo de gravação das informações é efetivado quando o usuário definitivamente escolhe os termos que gostaria de usar como *tags* e grava as informações.

Ao decretar definitivamente o conjunto de dados que descrevem o objeto digital, o sistema receberá e armazenará no banco de dados o conjunto de informações que o usuário escolheu para descrever o objeto digital.

O processo denominado Folksonomia Assistida efetiva a necessidade de fazer com que as *tags* tenham um grau maior de significado em relação ao objeto depositado, principalmente dentro do contexto em que está sendo utilizada. Guy e Tonkin (2006, p. 1) afirmam que,

Começamos por olhar para a questão das "tags malfeitas", um problema para o qual os críticos da Folksonomia fazem questão de aludir, e perguntar à comunidade que pesquisa sobre Folksonomia se há maneiras de compensar esses problemas [...] Provavelmente, a grande falha dos sistemas de folksonomia atuais, é que os termos de marcação utilizados nesses sistemas são imprecisos. Os usuários dos sistemas que utilizam Folksonomia inserem livremente as tags, o que significa que as tags são muitas vezes ambíguas, excessivamente personalizadas e inexatas.

O uso da Folksonomia Assistida busca justamente melhorar a eficiência do uso de *tags*, permitindo ao usuário uma descrição livre para os objetos digitais a que deposita, porém de forma que possa se amparar no próprio conhecimento já disponível no ambiente em que está utilizando.

Ressalta-se que no ano de 2012 houve uma homologação da processo de Folksonomia Assistida. Através do desenvolvimento de um modelo teórico-conceitual. Nesta homologação, utilizou-se um conjunto de palavras chaves de cinquenta artigos e fora possível verificar, entre outros resultados, uma diminuição de 12,79% das palavras-chave em relação ao depósito inicial. Os resultados homologam a vantagem de uso da Folksonomia Assistida, ao possibilitar a visualização da congregação dos relacionamentos entre as *tags* anteriormente dispersas pelas inconsistências. De forma cooperativa, desempenham

um papel valioso no gerenciamento do vocabulário utilizado na representação informacional, em aspectos que contribuem para a não redundância de informação, e na confiabilidade da pesquisa e recuperação dos artigos pelos usuários do repositório digital (SANTARÉM SEGUNDO; SIQUEIRA, 2013).

### Iteratividade, a retroalimentação da informação

A implementação do processo de Folksonomia Assistida é a base para a consolidação da Representação Iterativa, que deve ser retroalimentada, sempre baseada no contexto de uma estrutura de representação do conhecimento, através de uma ontologia, que consiste em definir os limites de um domínio do conhecimento.

Essa visão geral é detalhada na Figura 1 que apresenta os passos para que realmente aconteça o uso completo da Representação Iterativa.

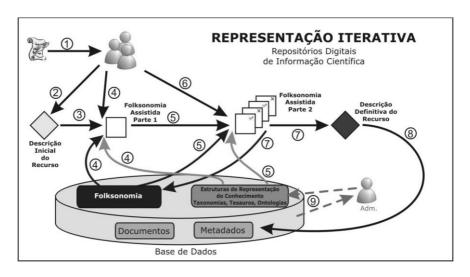


Figura 1: Representação Iterativa — Visão Detalhada Fonte: Santarém Segundo (2010)

A construção do modelo nomeado Representação Iterativa, apresentado de forma detalhada na Figura 1, deverá ser realizada conforme os seguintes passos:

- 1. Os usuários fazem uma leitura e verificação do documento a ser depositado e extraem os metadados necessários que descrevam o máximo possível o objeto, para que seja realizado o depósito;
- 2. O usuário através de formulário disponível no ambiente inicia o processo de descrição do recurso. Esse passo é chamado de descriçao inicial do recurso porque é neste momento em que o usuário deverá inserir todos os metadados relativos ao objeto, com exceção da *tag* assunto;
- 3. Com as informações dos metadados já alimentadas, o sistema encaminha o usuário para fazer a descrição da tag assunto, que é a informação que representará de forma mais significativa o recurso dentro da Representação Iterativa;
- 4. Esse passo representa o início da Folksonomia Assistida. Nesse momento, o sistema deverá colaborar na descrição da *tag* utilizando uma estrutura de sugestão semelhante ao da pesquisa do Google, sendo que as informações sugeridas serão os próprios termos já inseridos anteriormente por usuários (folksonomia representação livre), além dos termos que fazem parte da estrutura de representação do conhecimemento (taxonomias, ontologias ou tesauros) que estará associada ao repositório;
- 5. Nesse passo acontece o segundo momento da Folksonomia Assistida. Após a descrição da *tag* assunto, o ambiente reconhece essas informações e busca relacionamentos e associações dentro do instrumento de estrutura de representação do conhecimento utilizado, agrega termos, e em seguida faz o mesmo dentro do conjunto de *tags* já definidas por outros usuários (representação livre), busca associações e, na sequência, também relaciona termos.

Esse conjunto de termos que foram selecionados são devolvidos para o usuário;

- 6. O usuário volta a atuar novamente assim que recebe o conjunto de termos do ambiente. Neste momento ele deve completar o processo de Folksonomia Assistida, escolhendo de forma definitiva os termos que serão utilizados na *tag* assunto. Essa decisão implica em estabelecer relacionamentos entre as *tags*, e, portanto, criar a relação semântica de termos que irá caracterizar a recuperação semântica posterior. Portanto, esse momento é muito importante para a consolidação da Representação Iterativa, porque estabelece os termos e relacionamentos que caracterizam o recurso;
- 7. Esse passo apenas apresenta a confirmação da descrição completa do recurso, visto que o usuário já descreveu inicialmente os metadados e em seguida, com auxílio da Folksonomia Assistida, escolheu os termos que compõe a tag assunto. É nesse momento que a folksonomia será alimentada efetivamente com o novo conjunto de termos e relacionamentos que o usuário efetivou e dessa forma reorganizada, atualizando o peso dos termos e relacioanamentos de acordo com os novos elementos que foram inseridos. Cada vez que esse passo é efetivado em um novo depósito acontece um enriquecimento e fortalecimento do conjunto de termos e relações existentes, e as informações que foram inseridas passam a ficar disponíveis para serem utilizadas por novos usuários em novos depósitos;
- 8. Nesse passo o conjunto completo de metadados, assim como os objetos digitais, são armazenados na base de dados;
- 9. A cada período de tempo o processo deverá ser avaliado por um administrador de sistema que poderá também retroalimentar a estrutura de representação do conhecimento das áreas de especialidades, dando uma nova visão a respeito dos limites estabelecidos ao domínio do conhecimento. Esse processo cria uma

nova perspectiva na Ciência da Informação, que é a avaliação e reconstrução da estrutura de representação do conhecimento, baseado na construção da informação, por usuários de um ambiente digital.

É importante ressaltar que o administrador deve ser um profissional ou equipe multidisciplinar responsável pela catalogação do ambiente informacional e pela manutenção das estruturas de representação do conhecimento (bibliotecário, arguivísta e/ou cientista da informação).

O processo de iteratividade é estabelecido de forma que fique a cargo de um usuário administrador a retroalimentação da estrutura de representação do conhecimento, e, como função sistemática e automática dos usuários, as retroalimentações da folksonomia.

O processo de iteratividade resulta na reconstrução do conhecimento de forma coletiva e moderada, permitindo o enriquecimento e amadurecimento da estrutura de representação do conhecimento para o domínio em que o repositório digital científico está inserido.

A arquitetura proposta neste trabalho parte do princípio da iteratividade, que é o processo em que ocorre a realimentação constante do sistema em busca da melhor qualidade do conjunto de informações.

O princípio da iteratividade está dentro do contexto de desenvolvimento de software, do qual foi realizada uma adaptação para a construção deste modelo estrutural para repositórios digitais científicos. É importante ressaltar que o estudo de processos e metodologias para melhorar o desenvolvimento de software é constante dentro da área de Ciência da Computação.

Segundo Larman (2007, p. 47),

O ciclo de vida iterativo é baseado em refinamentos e incrementos sucessivos de um sistema por meio de múltiplas iterações, com realimentação (feedback) e adaptação cíclicas como principais propulsores para convergir para um sistema adequado. O sistema cresce incrementalmente ao longo do tempo, iteração por iteração, razão pela qual esta abordagem também é conhecida como desenvolvimento iterativo e incremental. Como a realimentação e adaptação fazem as especificações e o projeto evoluir, esse sistema é conhecido como desenvolvimento iterativo e evolutivo.

Larman (2007) afirma que o processo iterativo é também evolutivo, assim como acontece no modelo proposto nesta pesquisa, que propõe a evolução das representações de informação.

Essa evolução pode ocorrer através da Folksonomia Assistida, ou seja, do processo repetitivo de inserção de conteúdos para *tags*, assim como da evolução e adaptação da estrutura de representação do conhecimento utilizada, por intermédio de um administrador.

O processo de desenvolvimento iterativo, do qual a Representação Iterativa é baseada, deve obedecer limites temporais.

Larman (2007, p. 50) alerta que: "A maioria dos métodos iterativos recomenda que a duração de uma iteração seja entre duas e seis semanas".

A Representação Iterativa tem um contexto diferente, porque não trata de desenvolvimento de software, mas sim da construção do corpus de informação de um domínio, através de uma inteligência coletiva, porém o princípio da temporalidade também pode ser abordado e utilizado.

O processo de construção da inteligência coletiva pela Folksonomia Assistida não deve ser temporal. Ele deve ser feito dinamicamente, sugerindo que o usuário possa ter acesso ao conjunto de informações a qualquer momento, ou seja, assim que uma  $t\alpha g$  é inserida no sistema, ela passa a ficar disponível para ser utilizada como sugestão a outros usuários.

O acesso do administrador do sistema, para fazer ajustes ao modelo que está sendo construído, pode ter sim uma temporalidade definida, corroborando com a ideia de iteratividade. O modelo Representação Iterativa não define um intervalo exato de temporalidade de intervenção do administrador do sistema, porém cada ambiente deve estabelecer seu próprio intervalo de temporalidade de acordo com a quantidade de acessos e o volume de informações dentro do repositório (SANTAREM SEGUNDO, 2010).

Cada iteração gera um novo conjunto de informações, relacionamentos e também uma forma diferente de conhecimento. Dentro do contexto evolutivo do modelo, é possível que a interatividade entre os usuários e o sistema gere uma camada de informações cada vez mais rica, principalmente porque permite feedback ao usuário, assim como a possibilidade da informação já armazenada servir como base para que a próxima seja inserida.

Dessa forma, a Representação Iterativa oferece aos repositórios um novo formato de organização da informação, de modo que passe a existir uma relação entre os trabalhos autoarquivados, não apenas pela simples sintaxe das palavras-chave e nem tampouco pela comunidade e coleção de que fazem parte.

A estrutura funcional deste modelo parte do princípio da agregação de valores ao repositório, de forma que ocorra uma contextualização do material digital inserido, criando relações que possam sustentar uma recuperação semântica de informações.

# Recuperação da informação na Representação Iterativa

A estrutura da Representação Iterativa permite criar novos sistemas de recuperação da informação dentro dos repositórios. Os novos modelos não devem substituir o anterior, mas sim agregar mais opções de pesquisa e interação do usuário com o ambiente.

Um dos formatos propostos é o de rede, que tem sido muito abordado no conceito de colaboração científica, principalmente nos relacionamentos entre co-autorias e co-citações. Porém, a mesma ideia utilizada neste conceito se aplica às redes de *tags*, que podem agregar a informação a respeito dos autores e criar o conceito de autores que tenham o mesmo perfil de depósito dentro de um repositório. O modelo em formato de rede aproxima termos que estão relacionados, criando uma estrutura de informação que tem apresentação visual agradável e de entendimento intuitivo.

Portanto, se a ideia de co-autoria das redes colaborativas torna possível uma grande quantidade de estudos, as redes construídas por meio da estrutura da Representação Iterativa poderão gerar um conjunto grande de informações a respeito do conteúdo dos objetos depositados nos repositórios digitais informacionais.

Para a criação de uma rede de informações que permita ao usuário navegar pelos termos, os dados registrados deverão formar uma matriz de adjacência que possibilite a construção do gráfico, que é a estrutura matemática e computacional escolhida para representar as redes.

A apresentação do gráfico possibilita algumas variações. Neste trabalho, sugere-se que os vértices tenham tamanhos diferentes, de forma proporcional, de acordo com a quantidade de incidências dos termos armazenados no repositório.

A implementação do processo de recuperação ainda não está finalizado, porém a indicação é do uso das estruturas de gráfico e

linguagem Sparql para recuperação da informação através do uso de ontologias.

# Considerações

A construção desta pesquisa nasceu da necessidade de melhorar a recuperação da informação em repositórios digitais informacionais. Para que isso fosse possível, foi proposto um modelo novo nomeado Representação Iterativa, para repositórios digitais.

O modelo conceitual ainda está em processo de implementação e vem evoluindo conforme o tempo. Destaca-se que já é possível perceber novas pesquisas que se apoiam no modelo proposto para concepção de outros trabalhos com características parecidas.

Dessa maneira, entende-se que tanto o processo, já homologado, de Folksonomia Assistida, quanto todo o modelo nomeado Representação Iterativa, cumprem com o propósito de qualificar a produção e recuperação de informação em contextos semânticos aplicados em repositórios digitais.

Essas características apresentadas indicam claramente a possibilidade de apoio de estruturas como ontologias para a concepção de modelos semânticos em repositórios digitais, facilitando o processo de organização da informação e principalmente da recuperação desta através de relações semanticamente estruturadas.

# **REFERÊNCIAS**

AQUINO, M. C. *Hipertexto 2.o, folksonomia e memória coletiva*: um estudo das *tags* na organização da *web*. E-Compós, Brasília, v. 9, 2007. Disponível em:<a href="http://www.compos.org.br/seer/index.php/e-compos/article/view/165/166">http://www.compos.org.br/seer/index.php/e-compos/article/view/165/166</a>>.

Acesso em: 3 set. 2013.

BEKAERT, J.; VAN DE SOMPEL, H. *Augmenting interoperability across scholarly repositories*. Report, 2006. Disponível em: <a href="http://msc.mellon.org/Meetings/">http://msc.mellon.org/Meetings/</a> Interop/FinalReport >. Acesso em: 14 set. 2013.

CAFÉ, L. et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: CONGRESSO BRASILEIRO DE CIÊNCIAS DA COMUNICAÇÃO, 26. 2003, Belo Horizonte. *Anais...* Belo Horizonte: INTERCOM, 2003. Disponível em: <a href="http://www.intercom.org.br/papers/nacionais/2003/www/pdf/2003\_ENDOCOM\_TRABALHO\_cafe.pdf">http://www.intercom.org.br/papers/nacionais/2003/www/pdf/2003\_ENDOCOM\_TRABALHO\_cafe.pdf</a> >. Acesso em: 2 set. 2013.

GUY, M.; TONKIN, E. Folksonomies: tidying up tags? *D-Lib Magazine*, Reston, v.12, n.1, Jan. 2006. Disponível em: <a href="http://wwww.dlib.org/dlib/januaryo6/guy/o1quy.html">http://wwww.dlib.org/dlib/januaryo6/guy/o1quy.html</a>>. Acesso em: 15 set. 2013.

LARMAN, C. *Utilizando UML e padrões*: uma introdução à análise e ao projeto orientados a objetos e ao desenvolvimento iterativo. 3. ed. São Paulo: Bookman, 2007.

SANTAREM SEGUNDO, J. E. *Representação Iterativa*: um modelo para repositórios digitais. 2010. 224 f. Tese (Doutorado em Ciência da Informação) — Faculdade de Filosofia e Ciências, Universidade Estadual Paulista, Marília. 2010.

SANTAREM SEGUNDO, J. E.; DONADI, R. O profissional da informação no contexto das TICS's: aplicação do conceito de WIKI em empresas de desenvolvimento de software. In: CASTRO FILHO, C. M. (Org.). *Olhares sobre o profissional da Ciência da Informação*. São Paulo: Editora Todas as Musas; 2013. p. 163-184.

SANTAREM SEGUNDO, J. E.; SIQUEIRA, C. S. Aplicación teórico-conceptual de folksonomías asistidas para la recuperación de información. Scire, Zaragoza, v. 19, n. 2, p. 77-82, 2013.

WAL, T. V. Folksonomy definition and wikipedia. Disponível em: <a href="http://www.vanderwal.net/random/entrysel.php?blog=1750">http://www.vanderwal.net/random/entrysel.php?blog=1750</a>. Acesso em: 2 set. 2013.

# Bibliotecas e repositórios digitais: reflexões, tecnologias e aplicações

Divino Ignacio Ribeiro Jr1

# Introdução

É amplamente conhecido o fato que nos últimos sessenta anos a popularização das Tecnologias da Informação e Comunicação (TIC) trouxe muitas mudanças de ordem econômica, social e cultural. O paradigma de produção industrial e a riqueza medida em termos de bens materiais se reconfigurou com a emergência do conceito de Sociedade da Informação, e posteriormente, da Sociedade do Conhecimento.

Informação e Conhecimento passam a ser vistos como bens de consumo e insumos de produção, ofertados por serviços especializados providos por empresas e profissionais com formação multidisciplinar, caracterizando um novo tipo de mercado e de formação. Novos cursos de graduação e pós-graduação foram criados e/ou adaptados, áreas do conhecimento científico e tecnológico se consolidaram ao longo dos anos, tendo como eixo a evolução das TIC.

Tal paradigma socioeconômico se sustenta, inicialmente, sobre a popularização da infraestrutura de telecomunicações, com o registro

<sup>&</sup>lt;sup>1</sup> Professor do Programa de Pós-Graduação em Gestão da Informação da Universidade do Estado de Santa Catarina (UDESC), coordenador do Laboratório de Tecnologias Aplicadas à Gestão do Conhecimento e da Informação (LabTecGCI /UDESC). CONTATO: divinoirj@gmail.com

crescente de domicílios com computador e uso da Internet. Dados da pesquisa TIC Domicílios 2010 realizada pelo Comitê Gestor da internet no Brasil nos mostram que cada ano mais e mais pessoas têm acesso à internet e computador em suas residências, e tal tendência se verifica até os dias atuais. Vejamos os dados a seguir:

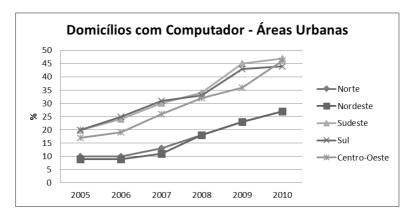


Figura 1 - Crescimento do % de Domicílios com Computador no Brasil (fonte: Cetic.BR)

Outro importante indicador que nos mostra essa popularização é o crescimento do uso da internet nos domicílios brasileiros:

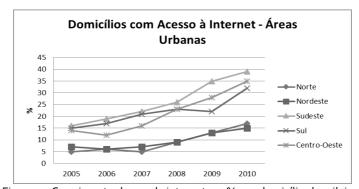


Figura 2 - Crescimento do uso da internet em % nos domicílios brasileiros

Tais indicadores são evidências de que cada vez mais a população tem acesso aos meios de comunicação (especificamente uso de computador e internet), o que proporciona novas possibilidades para a disseminação da cultura digital.

Se esses indicadores<sup>2</sup> seguirem essa tendência, em cerca de 10 anos a maior parte da população do Brasil terá acesso à internet e computador em suas casas.

Nessa linha de pensamento, Sayão *et al.* (2009) afirmam que o século XXI consolida a cultura da disseminação da informação eletrônica; por meio da internet, o acesso às fontes de informação e canais de comunicação, como listas de discussão, *blogs*, *chats*, mídias sociais, entre outros.

Essa cultura da disseminação da informação eletrônica movimenta a indústria de equipamentos de informática, telecomunicação e outros, que utilizam tecnologias com fins computacionais e para comunicação em rede; impulsiona o mercado de serviços de telefonia móvel, acesso à internet, serviços de telefonia VoIP – *Voice over IP*, sistemas de IPTV – transmissão de sinal de TV através de canais de dados, entre outros setores.

Os investimentos em produtos e serviços dependentes dessas tecnologias são cada vez maiores: por exemplo, as empresas de telecomunicações modificam suas redes analógicas para redes digitais de dados (ex: substituição de cabeamento metálico por cabeamento de fibra ótica) e modificam as tecnologias para funcionamento através delas; empresas de TV a cabo estão transmitindo sinal digital em sua rede, oferecendo através do mesmo meio serviços de TV, Dados e Voz.

Há fortes investimentos na 'geração de demanda' para que a população continue aderindo a novos serviços; as empresas de telecomunicação,

\_

<sup>&</sup>lt;sup>2</sup> O portal CETIC.BR publicou a série de indicadores 2013, disponível em www.cetic.br

por exemplo, não vendem 'celulares' ou 'planos de telefonia e dados' – todo esforço de marketing é focado para vender 'conectividade', formar e alimentar a necessidade de estar sempre 'conectado' e 'acessível'. A partir daí instala-se um processo retroalimentado: a população desenvolve essa necessidade de consumo de serviços de telecomunicação, impulsionada pelas empresas que os fornecem. Outros mercados, como o de aplicativos para telefonia móvel, oferecem serviços de *software* em vários segmentos, criando novas demandas de consumo.

Além do aspecto econômico, há as dimensões sociais e culturais. Os impactos da cultura digital permeiam a formação de novas gerações, influenciando a maneira como pensam, escrevem, se comunicam.

O meio científico e acadêmico não é exceção: ao longo do tempo novos produtos de informação surgiram para responder a este cenário de mudanças constantes. No próximo tópico trataremos dos Repositórios Digitais e das Bibliotecas Digitais, exemplos incluídos nesse contexto.

# Repositórios digitais

O surgimento dos Repositórios Institucionais é, em última análise, o resultado de um processo de mudanças no processo de comunicação científica ocorrido nos últimos anos.

A história do surgimento dos Repositórios Institucionais tem relação com o crescimento do mercado editorial de periódicos científicos. Tal crescimento se deu, fundamentalmente, pelo constante crescimento dos investimentos na atividade científica ocorrida após a Segunda Guerra Mundial e o consequente surgimento de periódicos científicos especializados nas mais diversas áreas do conhecimento.

No entanto, a legitimidade e o valor dos trabalhos científicos estão ligados diretamente à visibilidade proporcionada pelos canais de comunicação científica; a consolidação do mercado editorial como um

modelo de negócio altamente lucrativo provocou distorções nesse processo de comunicação.

Em princípio, esses canais deveriam levar à comunidade acadêmica e à população em geral os resultados da atividade científica, que em muitas situações era financiada com recursos governamentais. O acesso a esses resultados, necessário à continuidade e à evolução de novas pesquisas só era possível por parte de quem pudesse pagar os altos custos das assinaturas, restringindo assim a comunicação científica para instituições com recursos suficientes para custeá-las, excluindo a comunidade acadêmica que não tivesse recursos financeiros para pagar pelo acesso.

Além da dificuldade de acesso, esse modelo de negócio do mercado editorial instalou um círculo vicioso: a obrigação de publicar em revistas científicas reconhecidas pelas comunidades acadêmicas, já absorvidas por grandes grupos desse mercado editorial; as métricas de impacto e de produtividade consideram de maneira preponderante as revistas que pertencem a esse círculo restrito ao mercado editorial.

Tal situação culminou no que a literatura chama de "crise do periódico científico" no final dos anos 90 o que motivou, segundo Kuramoto (2006) o seguinte fato:

Em julho de 1999, Paul Ginsparg, Rick Luce e Herbert Van de Sompel fizeram uma chamada para uma reunião exploratória entre os responsáveis por repositórios de eprints acadêmicos, a Convenção de Santa Fé. A partir desse evento, foi criada a Open Archives Initiative (OAI). A meta principal dessa iniciativa é contribuir de forma concentrada para a transformação da comunicação científica. A linha de ação proposta para essa transformação é a definição de aspectos técnicos e de suporte organizacional de uma estrutura de publicação científica aberta, na qual ambas, a camada comercial e livre, possam se estabelecer.

Esse foi o marco inicial para o que se conhece atualmente como **Iniciativa de Arquivos Abertos (Open Archives Initiative – OAI)**. Outros marcos importantes ocorreram:

- 1999 Convenção de Santa Fé
- 1999 Proposta PubMed Central (PMC)
- 2001 Carta aberta da Public Library of Science (PloS)
- 2002 Carta de Budapeste para o Acesso Aberto
- 2003 Declaração de Bethesda
- 2003 Declaração de Berlim
- 2003 Declaração da IFLA para o Acesso Aberto
- 2004 Resolução da câmara dos comuns, Reino Unido
- 2004 Declaração de Valparaiso
- 2004 Publicação do relatório do comitê britânico
- 2005 Manifesto de acesso livre à informação em C&T IBICT
- 2005 Declaração de Salvador

A partir de então se verifica a realização de diversas ações para consolidação de tecnologias, trocas de experiências, realização de eventos (*Workshops*, Seminários, entre outros) dirigidas à promoção e criação de Repositórios Institucionais.

Um documento que sintetiza com maestria essa iniciativa é o *Manifesto Brasileiro de apoio ao Acesso Livre à Informação Científica* lançado pelo IBICT em 2005, seguindo uma tendência mundial no apoio à Iniciativa de Arquivos Abertos (*Open Archives Initiative* – OAI).

É a partir dele que Sayão et al. (2009) destacam um trecho que caracteriza, fundamentalmente, o papel dos Repositórios Institucionais:

Uma versão completa da obra e todos os materiais suplementares, incluindo uma cópia da licença, como acima definida, é depositada e, portanto, publicada em um formato eletrônico normalizado e apropriado em pelo menos um repositório que utilize normas técnicas adequadas (como as

definições estabelecidas pelo modelo Open Archives) e que seja mantido por uma instituição acadêmica, sociedade científica, organismo governamental, ou outra organização estabelecida que pretenda promover o acesso livre, a distribuição irrestrita, a interoperabilidade e o arquivamento a longo prazo.

Podemos interpretar essa transcrição do Manifesto como uma forma abrangente de conceito sobre repositórios; se atentarmos para a literatura da área encontraremos diversos autores que apresentam conceitos similares, como por exemplo:

# Repositório Temático:

Conjunto de trabalhos de pesquisa de uma determinada área do conhecimento disponibilizados na internet. Esses repositórios utilizam tecnologias abertas e seguem a filosofia da Iniciativa dos Arquivos Abertos, promovendo a maior acessibilidade à produção dos pesquisadores e à discussão entre seus pares - Café et al. (2003);

# Repositório Institucional:

Um repositório institucional é a reunião de todos os repositórios temáticos hospedados em uma organização. No caso de uma universidade, cada departamento trata de uma área do conhecimento e, portanto, seu repositório temático será específico no assunto deste departamento. A união de todos os repositórios das diversas unidades de pesquisa comporá o repositório institucional, caracterizando-o como multidisciplinar - Café *et al.* (2003);

Englobam a produção científica de determinada instituição, mais comumente institutos de pesquisa e universidades. Hospedam, geralmente, uma coleção de documentos de pesquisa (pré-prints e pósprints), embora possam incluir relatórios técnicos, manuscritos, dados, videoclipes e imagens, além de conter dados administrativos de apoio à instituição, como arquivo local de documentação, teses, dissertações, livros e outros - Boso (2011);

### Repositório Acadêmico:

Um serviço cuja ênfase é constituir, gerenciar e sobretudo, disseminar coleções digitais de informação científica de modo que a comunicação, acesso e uso de resultados de pesquisa fossem expandidos. Esse serviço possui as seguintes características: a) é institucionalmente definido; b) é academicamente orientado; c) é cumulativo e perpétuo e d) é aberto e interoperável (COSTA e LEITE, 2009).

Por fim, o conceito de Repositório tem sido adaptado por conta de novos usos desse tipo de produto de informação. Há dezenas de exemplos de Repositórios destinados à publicação de acervos e obras literárias completas (ex: Brasiliana Digital³) e outros tipos de acervos digitais, como materiais em áudio e vídeo (ex: Instituto Paulo Freire⁴).

Tais repositórios não têm o mesmo escopo e finalidade daqueles criados em âmbito acadêmico, mas possuem as mesmas características essenciais do que podemos denominar simplesmente <u>Repositório Digital</u>.

Com base nas definições da literatura e nas experiências institucionais, podemos definir um produto de informação como Repositório Digital quando ele possuir as seguintes características:

- Ser uma versão completa da obra e todos os materiais suplementares, incluindo uma cópia da licença, depositada com o material;
- Publicada com padrões tecnológicos aderentes a normas técnicas de preservação digital (como as definições estabelecidas pelo modelo Open Archives e o modelo OAIS);
- Mantido por uma instituição acadêmica, sociedade científica, organismo governamental, setor privado, ou outra organização

<sup>&</sup>lt;sup>3</sup> Visite o site da Brasiliana Digital em http://www.brasiliana.usp.br/bbd

<sup>&</sup>lt;sup>4</sup> Visite o Centro de Referência Paulo Freire: http://acervo.paulofreire.org

estabelecida que pretenda promover o acesso, a distribuição, a interoperabilidade e o arquivamento em longo prazo.

# **Bibliotecas digitais**

O conceito de Biblioteca Digital evoluiu bastante ao longo dos anos. Influenciado pela inserção das TIC nas Bibliotecas, esse conceito ganhou vários matizes. É possível observar na literatura muitas especulações, reflexões e estudos acerca das diferentes dimensões de uma biblioteca: da gestão de acervos, de pessoas, dos processos administrativos, dos serviços oferecidos, entre outros.

Drabenstott e Burman (1997) nos contam que concepções de uma 'biblioteca do futuro 'ou uma 'biblioteca inteligente' estão registradas na literatura desde os anos 6o. Os autores citam o trabalho de J. C. R. Llicklider, intitulado Libraries of the future, visualizando um cenário avançado de busca de documentos legíveis por computador, acessando simultaneamente várias bases de dados.

Na época em que esse artigo foi publicado pelas autoras, elas identificaram 15 definições sobre bibliotecas digitais, entre as quais destacam: biblioteca digital, biblioteca eletrônica, biblioteca virtual, biblioteca biônica, biblioteca sem paredes, biblioteca do futuro, máquina eletrônica.

Elas enumeram os recursos oferecidos por esse tipo de biblioteca (preservamos as ideias com a mesma terminologia da época):

- Acesso a coleções de diversas bibliotecas para bases de dados e outras fontes;
- Livros virtuais;
- Redes de catálogos on-line;
- Conexão de qualquer lugar, por meio de equipamentos especializados;
- Disponibilidade 'infinita' de coleções;

- Acesso e transferência de conteúdos 'quantas vezes forem necessárias, sem constrangimento';
- Acesso a arquivos na íntegra, imagens digitais, audiovisuais, animações, sem precisar sair de casa, via modem e tecnologias eletrônicas.

Um caso bastante citado na literatura dos anos 90 é o da empresa de cosméticos Natura, que, naquela ocasião, adotou a concepção de 'acesso' em detrimento do 'acervo', iniciando uma política de acessos e intercâmbios com entidades detentoras de informação, em nível nacional e internacional, principalmente via banco de dados *on-line*. Ao criar sua biblioteca, concebeu um organismo que tem como papel principal fornecer informação ao invés de documentos; preservar a memória dos processos e criação dos produtos ao invés de acumular exemplares de livros (REZENDE, 1997).

Em 1999 Murilo Bastos da Cunha publica um artigo intitulado Desafios na construção de *uma biblioteca digital*, onde trata de temas sobre a possibilidade de uma biblioteca digital, seus paradigmas, apresentando questões como: a) ainda existirão livros no futuro? b) deve-se continuar a assinar periódicos impressos, em CD-ROM ou aguardar o periódico totalmente eletrônico?

O autor aborda em seu texto os seguintes aspectos que devem ser considerados na construção de uma biblioteca digital:

- Instalações Físicas: naquela época as bibliotecas não eram construídas para acomodar a infraestrutura tecnológica, que então, requeria grandes espaços para funcionar e um processo de adaptação de custo alto;
- Aquisição, desenvolvimento de coleções e comutação bibliográfica: com a inserção dos computadores, redes de comunicação e dos documentos digitais, redefinem-se todas questões relativas ao desenvolvimento de coleções e gestão de estogues informacionais;

- Catalogação, classificação e indexação: acervos digitais requerem tratamento diferenciado, em alguns aspectos, dos documentos físicos. Por exemplo, faria sentido usar algum sistema de classificação para localizar o documento, quando na realidade, esses documentos são localizados por meio de uma interface de busca?
- Periódicos: naquela época os sistemas de construção de periódicos eletrônicos não faziam parte do dia a dia das bibliotecas. Segundo o autor, em 1998 no Brasil, haviam 133 títulos de periódicos eletrônicos, construídos com tecnologias baseadas em html. O futuro dos periódicos nesse contexto ainda era uma questão muito nebulosa, segundo o autor;
- Serviço de Referência: o bibliotecário de referência deixará de existir? Será substituído por um 'serviço de referência eletrônica'?
- Preservação: como preservar os documentos digitais e seus suportes?
- Tecnologias: como lidar com a obsolescência da tecnologia da informação?

Algumas dessas questões, ainda que antigas, são pertinentes; o enfoque ou o tipo de pergunta no escopo de cada tema mudou ao longo dos anos, mas alguns desses temas ainda estão presentes quando se trata da criação de bibliotecas digitais.

Quase dez anos depois, o mesmo autor resgata a temática em seu artigo *Das bibliotecas convencionais às digitais: diferenças e convergências* (Cunha, 2008), com uma discussão mais concreta do que as que relatamos neste texto.

# Ele aponta as seguintes diferenças:

 A organização da Informação: o tratamento documentário, o armazenamento e a recuperação guardam diferenças substanciais, mas os métodos e processos das bibliotecas convencionais não desapareceram, pelo contrário, foram adaptados para o mundo digital; na convencional, 'o usuário vai à biblioteca', na digital, a 'biblioteca vai ao usuário;

- O acesso à informação: o paradigma da integração de fontes de informação é um requisito indispensável;
- O aspecto econômico: na biblioteca digital, é possível, por meio da integração, dispor acesso a um número grande de fontes, com custos relativamente reduzidos em relação à biblioteca convencional;
- O incremento da cooperação de serviços entre as bibliotecas, uma vez que as tecnologias e padrões facilitaram a comunicação entre diferentes plataformas de maneira rápida e com baixo custo.

As mudanças sobre a concepção das Bibliotecas Digitais ocorrem no mundo todo. Isso pode ser verificado a partir das discussões realizadas pela IFLA - *International Federation of Library Associations*<sup>5</sup>, uma instituição internacional que representa os interesses de bibliotecas e serviços de informação e seus usuários, na condição de um agente global que representa os profissionais ligados à Biblioteconomia e Ciência da Informação.

No Manifesto da IFLA para Bibliotecas Digitais encontramos a importância, o papel, missão, objetivos e recomendações para a criação dessas bibliotecas. Vejamos alguns trechos do Manifesto:

Atenuar a exclusão digital é um fator fundamental para alcançar os Objetivos de Desenvolvimento do Milênio das Nações Unidas. O acesso a recursos de informação e aos meios de comunicação apoia a saúde e a educação tanto quanto a cultura e desenvolvimento econômico. [grifo nosso] ... as bibliotecas operam digitalmente e os seus serviços digitais tornam acessível um novo canal para o universo de conhecimento e da informação, conectando

\_

<sup>&</sup>lt;sup>5</sup> Saiba mais sobre a IFLA em http://www.ifla.org/about/more

**culturas através de fronteiras geográficas e sociais.** [grifo nosso]

. . .

A missão da biblioteca digital é dar acesso direto a recursos de informação, digital e não digital, de forma estruturada e autorizada e, assim, ser uma ligação de tecnologia da informação, educação e cultura no serviço de uma biblioteca contemporânea [grifo nosso].

Para cumprir essa missão buscam-se os sequintes objetivos:

- Apoiar a digitalização, acesso e preservação do patrimônio cultural e científico.
- Proporcionar acesso a todos os usuários aos recursos de informação recolhidos pelas bibliotecas, respeitando os direitos de propriedade intelectual.
- Criar sistemas interoperáveis de biblioteca digital para promover padrões de livre acesso.
- Apoiar o papel essencial das bibliotecas e serviços de informação na promoção de normas comuns e as melhores práticas.
- Criar a consciência da necessidade urgente de garantir a acessibilidade permanente do material digital.
- Ligar as bibliotecas digitais de pesquisa de alta velocidade e redes de desenvolvimento.
- Aproveitar-se da maior convergência de meios de comunicação e papéis institucionais para criar e disseminar conteúdo digital.

A história das Bibliotecas Digitais possui uma trajetória permeada por questionamentos e experiências. É possível observar a evolução e a distinção de produtos e conceitos, juntamente com a evolução das TIC, sua inserção na sociedade e amadurecimento como modelos de negócio.

No âmbito acadêmico, um dos focos de discussão está na preocupação em se delimitar 'o que é' e 'quais os papéis' de uma biblioteca digital (e

no caso dos repositórios também). É uma discussão necessária, mas não suficiente, para subsidiar a criação de tais bibliotecas.

É possível verificar na literatura recente que uma Biblioteca Digital não se reduz às suas tecnologias; Biblioteca Digital não é um *software*, nem somente um acervo digital. Ela é, antes de tudo, uma Biblioteca, e como tal, possui caráter institucional, missão, valores, serviços, processos de gestão e avaliação, pessoas, infraestrutura.

O documento da IFLA é um bom exemplo da evolução dessa discussão: a ideia de Biblioteca Digital é definida de tal maneira que verificamos a maturidade e a convergência das experiências com a Gestão de Bibliotecas Digitais ao longo do tempo.

No Brasil, mesmo com iniciativas de pesquisa e formação de profissionais da área de Ciência da Informação, ainda há alguma carência de estudos objetivos que contemplem as dimensões de Gestão e de Tecnologias da Informação, de tal maneira que equipes empreendedoras de projetos de Bibliotecas Digitais tenham melhores condições para desenvolvê-los.

Outro fato a ser observado é que as Bibliotecas convivem com a 'tradição' – preservação de acervos físicos e a 'atualidade' – gestão da informação; seus processos e produtos transitam entre esses dois paradigmas. Miranda (1993) já havia constatado essa dualidade há cerca de vinte anos, ao afirmar que as bibliotecas estão se transformando de organizações centradas em coleções físicas para organizações centradas no acesso à informação.

Para ilustrar essa dualidade de paradigmas, vejamos a ilustração a seguir:

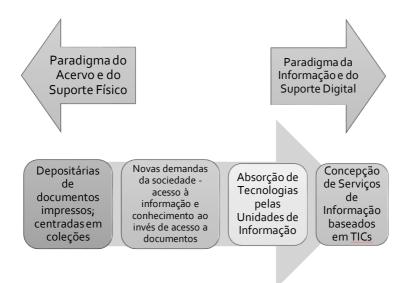


Figura 3 - Diferentes paradigmas que orientam serviços nas Bibliotecas

# Evolução tecnológica: o fio condutor das mudanças das Unidades de Informação

Ao longo do século passado, as Unidades de Informação, em especial as bibliotecas, sofreram mudanças intensas com os impactos da absorção das Tecnologias da Informação e Comunicação (TIC), com ênfase no contexto de seus processos internos e da criação de serviços. A evolução das TIC também provocou a criação de novos suportes de informação e de novos canais de comunicação, que igualmente provocaram mudanças nos processos e serviços das Bibliotecas.

Uma das razões que podemos considerar é o fato de que as mudanças organizacionais no contexto das unidades de informação, relativas à absorção das tecnologias da informação e comunicação, não ocorrem de forma brusca ou repentina, em outras palavras, tais mudanças não ocorrem em grandes saltos.

Isso pode ser verificado com uma rápida análise da história da absorção das tecnologias da informação e comunicação; vejamos a ilustração a seguir:

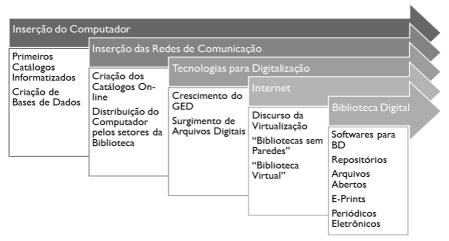


Figura 4 - Visão geral da absorção das Tecnologias da Informação e Comunicação pelas Unidades de Informação

# Inserção do computador

Os primeiros relatos de uso do computador nas unidades de informação datam da mesma época em que os computadores passaram a integrar a rotina das organizações.

O modelo de computação vigente nos anos 50 e 60 era predominantemente centralizado, com pouca ou nenhuma interação com os usuários dos setores da organização e sem uso de interfaces por meio de vídeo.

As primeiras experiências de informatização de catálogos em bibliotecas eram realizadas com apoio de processamento de dados para gerar listas impressas, fichas geradas por computador e índices classificados por computador.

Ao longo dos anos 60 foram desenvolvidas diversas tecnologias de equipamentos eletromecânicos para recuperação da informação, ou melhor, para apoiar a recuperação de documentos.

Tais equipamentos realizavam funções como a preparação de cartões perfurados nos quais eram registrados os índices para busca, máquinas classificadoras de cartões, máquinas que realizavam a localização (busca) e serviços análogos<sup>6</sup>.

Para muitos, imaginar como esses equipamentos funcionam é praticamente 'uma viagem no tempo', revivendo paradigmas das tecnologias da informação e da computação que há muitos anos não mais citados.

Com o surgimento dos terminais de vídeo e a descentralização gerada pela expansão do computador, pela organização, tais ferramentas começaram a incorporar as primeiras inciativas de interfaces de recuperação da informação.

# A Inserção das redes de comunicação

Um importante passo foi a popularização das redes de computadores. Paralelamente à diminuição de tamanho daqueles enormes computadores e do aumento da capacidade de processamento computacional, essa época foi marcada pelo surgimento dos catálogos *on-line* que permitiam a consulta às bases de dados via terminais dentro da organização.

As tecnologias computacionais para comunicação já permitiam a disseminação de terminais ligados a um computador central; esses terminais eram compostos por um monitor de vídeo e um teclado, por meio do qual o operador podia enviar solicitações de informação ao

<sup>&</sup>lt;sup>6</sup> Quer saber mais sobre essas tecnologias? Um livro excelente, porém de acesso relativamente difícil, é de Allen Kent, Manual da Recuperação Mecânica da Informação. Tradução de BERRINI, B. São Paulo: Polígono, 1972. 427 p.

computador central e obter respostas produzidas pelo processamento de dados.

A evolução e a miniaturização dos circuitos eletrônicos possibilitaram a criação dos mini e microcomputadores, e a consequente popularização do que se denominou na época como *Personal Computer – PC*.

Isso potencializou a disseminação das Redes de Computadores nas organizações, mudando radicalmente o conceito e o papel da informatização nas organizações.

A mudança mais perceptível nas bibliotecas foi a popularização das bases de dados via CD-ROM e *on-line*, e com recursos que permitiam ao usuário final realizar suas buscas em terminais dentro e fora da biblioteca.

# As tecnologias para digitalização

Outro importante marco foi a popularização das tecnologias para digitalização, verificada especialmente nos anos 8o. A facilidade de obtenção de *scanners* forçou a redução dos custos de aquisição e impulsionou um segmento de mercado muito importante para o contexto das unidades de informação de uma forma geral: GED - Gerenciamento Eletrônico de Documentos, ou mais recentemente, ECM – *Enterprise Content Managment*.

Por muitos anos uma das principais dificuldades da popularização da infraestrutura para GED foi o custo dos dispositivos de armazenamento de dados. O custo por *megabyte* era elevado, especialmente para o armazenamento magnético (discos rígidos) e os recursos para armazenamento óptico (CD-ROM e posteriormente em DVD) eram limitados e de difícil conservação.

Hoje em dia as tecnologias de scanners evoluíram bastante, há uma enorme variedade de tipos de tecnologias de hardware e software,

empresas especializadas em oferecer serviços de digitalização com uma relação custo-benefício muito atraente para as organizações.

# A Internet e as tecnologias para bibliotecas e repositórios digitais

A concepção tecnológica das Bibliotecas e Repositórios Digitais da atualidade só foi possível a partir da evolução dos padrões de engenharia de software para Web. O surgimento de novas linguagens de programação, softwares especializados em aplicativos com arquiteturas desenvolvidas para funcionar de maneira distribuída pela internet, constituem a base para as tecnologias, e consequentemente, para existência do que entendemos atualmente como Biblioteca Digital.

Como foi apresentado anteriormente, esse 'fio condutor' é um fator importante para redefinição dos paradigmas das Bibliotecas. Dessa maneira, podemos afirmar que essas instituições convivem com o 'passado' e com o 'futuro', lidando com processos de gestão e de serviços que atendem.

# Tecnologias da informação aplicadas em bibliotecas digitais

As bibliotecas e repositórios digitais só existem a partir das Tecnologias da Informação e Comunicação – TIC, e da sua popularização. Ao longo dos anos as unidades de informação têm absorvido as TIC para seu contexto organizacional e isso fez com que houvesse diversas mudanças em âmbito interno, nos processos de gestão e na maneira com que essas unidades se relacionam com seus usuários.

Os avanços tecnológicos influenciam cada vez mais as bibliotecas e bibliotecários, impondo uma demanda para absorção destes avanços para o alinhamento com as demandas informacionais e funcionais dos usuários de bibliotecas.

# O valor das tecnologias da informação e comunicação

Essa expressão – Tecnologias da Informação e Comunicação – é conhecida de muita gente. Largamente utilizada, refere-se a todo gênero de tecnologia aplicada ao armazenamento, processamento, recuperação e à comunicação de dados e informações.

Essas tecnologias são a base do mundo como conhecemos atualmente. Sem elas, muitos fenômenos, serviços, produtos das mais diversas ordens (individuais, organizacionais, sociais e econômicos) não ocorreriam ou sequer existiriam.

Como defini-las? Veja o quadro a seguir:



São recursos (insumos para processos da organização) baseados em tecnologias de computação e de comunicação em rede com capacidade para:

- Coletar e armazenar dados;
- Processar dados e produzir informação a partir deles;
- Recuperar dados e informações necessários aos processos operacionais e aos de gestão;
- Conectar pessoas e setores da organização;
- Criar canais de comunicação intraorganizacionais e com o ambiente externo à organização

Assim, as TIC têm papéis fundamentais no contexto organizacional e social, pois muitas vezes são literalmente a infraestrutura de diversos serviços, conceitos sociais, organizações, estilos de vida, entre outros.

Limitando esse conceito ao nível organizacional, temos alguns papéis fundamentais das TIC, na forma a seguir:

- Automação: aumentar a capacidade de auto-regulagem, auto-correção, auto-execução de processos ou eventos;
- Informatização: traduzir eventos e processos em dados, e submetê-los ao processamento para transformação em outros dados ou informações, com valor agregado;
- Recuperação da Informação: prover acesso a documentos e informação
- Comunicação: criar canais de comunicação com mecanismos de controle e gestão;
- Serviços baseados em Informação: criar novos serviços baseados em plataformas tecnológicas.

No entanto, as TIC têm diferentes participações em diferentes níveis da organização. Em outras palavras, elas estão presentes nos diferentes níveis organizacionais, adequadas e/ou adaptadas aos fins a que se prestam na organização. Vejamos a ilustração a seguir:

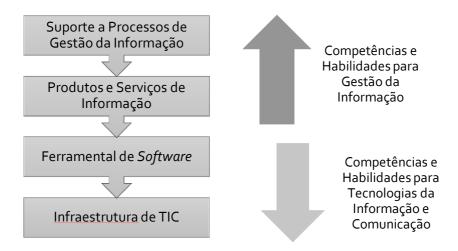


Figura 5 - Relação entre os níveis de TIC e perfis profissionais

A importância do papel das TIC depende do contexto em que ela está e do tipo de competências requeridas para sua gestão e utilização.

No nível mais essencial, o de Infraesturutura de TIC, estão: equipamentos de interconexão de rede, cabeamento estruturado, redes sem fio, servidores de rede (virtualizados ou não), computadores que funcionam como estações de trabalho, equipamentos especializados como scanners, estações de autoatendimento, equipamentos de apoio à identificação (leitoras, RFID), entre outros. Junto a essa infraestrutura estão pessoas com competências e habilidades para manter esses recursos operacionais, para gerenciar seu funcionamento, segurança e acesso.

No nível **Ferramental de** *Software* encontramos os recursos de apoio à gestão e uso dos recursos de Infraestrutura de TIC, como as plataformas de rede (sistemas operacionais de rede, diretórios, gerência de interconexão, *firewall*, sistemas para virtualização de servidores, serviço de VoIP, de videoconferência, sistemas de mensagens), e também, encontramos recursos para gerência de atividades de apoio às atividades fim: ferramentas de escritório, Sistemas de Gerenciamento de Bases de Dados – SGBD, *softwares* para apoiar a gestão de projetos, ferramentas de *workflow*, entre outros recursos que apoiem a criação de produtos e serviços da organização.

No nível **Produtos e Serviços de Informação** encontramos softwares de importância mais próxima às atividades-fim da organização, dependentes dos recursos das camadas anteriores, e ligados diretamente às atividades finalísticas da organização: *Softwares* para Repositórios Digitais, ambientes colaborativos (*wikis*, blogs), ferramentas de GED e ECM voltadas para o cliente final, ferramentas para Gestão de Bibliotecas (acervo, circulação, aquisição, etc.), gestão de arquivos físicos e digitais, entre outros.

No nível **Suporte a Processos de Gestão da Informação** encontramos os recursos que dão apoio direto ao processo decisório, ao

planejamento e gestão de políticas da organização. Tecnologias comumente envolvidas neste nível são recursos para Inteligência Competitiva, *Business Intelligence, Datawarehouse*, recursos gerenciais de *Enterprise resource planning* - ERP e outros recursos tecnológicos pertinentes a processos diretos de Gestão da Informação. Vale a pena ressaltar que aqui não se incluem tecnologias para Gestão do Conhecimento, que possuem outra categoria de tecnologias e outras finalidades.

Nesses contextos, a informação é um elemento chave e estratégico para realização de atividades, tanto aquelas do dia a dia quanto as que são de alta complexidade, sem a qual não podemos realizar ou atingir objetivos.

Podemos compreender, com base em nossas experiências, o valor que a informação possui no contexto de realização de nossas atividades. De acordo com Ribeiro Jr. (2001) não se trata de valor da mesma natureza que o bem de consumo tangível possui, mas de um valor agregado, que é estratégico, temporal e especializado.

- Estratégico: porque se caracteriza como um elemento que pode direcionar ou desencadear novos processos, seja em um negócio empresarial, seja para um trabalho escolar. Ao se administrar recursos, necessitamos de informação para que exista algum tipo ou grau de controle.
- **Temporal**: está relacionada ao contexto e ao tempo de vida útil que o valor de uma informação possui. Vale ressaltar que, relacionamos vida útil com a necessidade da informação e não com a informação em si mesma.
- **Especializado**: de forma análoga à temporalidade, também está relacionada à necessidade que se tem dela, ou de outra forma, se é ou não aplicável em um determinado contexto de necessidades.

Uma das questões mais complexas relacionadas à gestão da informação, de uma maneira geral, é estabelecer a importância que uma dada TIC possui em um contexto organizacional.

Uma TIC possui seu valor enquanto recurso potencialmente agregador de valor a processos, no entanto, esse não é o único, nem o principal parâmetro que define o seu valor.

Estabelecer o impacto final da aplicação de TIC no contexto de Bibliotecas Digitais segue o mesmo princípio; é necessário conhecer como elas funcionam, seus benefícios e custos (custo não se resume à questão financeira para sua aquisição, mas à sua manutenção e absorção pela organização).

Existe uma crença comum a muitos profissionais da informação de que as TIC sempre trazem benefícios, que a sua agregação é sempre boa para a organização. Isso cria uma visão distorcida em relação aos reais impactos da adoção de uma TIC em um dado contexto organizacional, acarretando prejuízos e problemas de ordem operacional e de gestão. Pelas mesmas razões (por essa crença equivocada) acredita-se que esses prejuízos foram trazidos pelas TIC adotadas, quando na verdade, o prejuízo operacional e de gestão foi causado por um processo decisório equivocado que resultou no mau dimensionamento de requisitos para a incorporação das TIC.

Alguns fatores que precisam ser considerados ao avaliar a importância de uma TIC num contexto organizacional:

 Quais os tipos de benefícios que serão agregados aos processos organizacionais nos quais uma TIC será inserida? Por exemplo, redução de tempo de execução de processo, aumento da confiabilidade dos resultados, produção de informação de alto valor agregado para processo decisório, produção de informação para auditoria, agregação de

- mecanismos de controle do fluxo informacional, entre outros;
- Qual é a relação de custo/benefício atual para obtenção ou produção de determinadas informações, considerando os processos nos quais uma TIC será inserida? Responder a essa questão auxilia na decisão sobre adotar ou não determinado recurso de TIC;
- Os parâmetros de avaliação da qualidade de uma dada TIC estão claramente definidos?
- Os serviços ou produtos que dependem diretamente do bom funcionamento das TIC toleram falhas, como por exemplo, paradas ou indisponibilidade temporárias?
- A organização consegue absorver o conhecimento tecnológico necessário à gestão das TIC que serão adotadas, para que sua manutenção e futura migração sejam sustentáveis?
- Há um plano de investimentos em infraestrutura de TIC, de maneira que seja possível avaliar a longevidade de um projeto (ex: de um repositório ou biblioteca digital) de serviço ou produto de informação?
- Qual é o comprometimento das instâncias de gestão da organização com investimentos para capacitação de recursos humanos para absorção do conhecimento tecnológico e no investimento para aquisição e atualização de TIC?

Esses, entre outros fatores, precisam ser questionados no momento em que planejamos a escolha de um recurso de TIC, seja ele qual for.

# Interoperabilidade: conceito chave para bibliotecas e repositórios digitais

Por quais razões deve-se optar pela interoperabilidade de sistemas? Podemos enumerar as seguintes razões:

## "Inter ⇔ operabilidade"

- Presume a operação com mútuas trocas de dados;
- Possibilidade de Integração de dados no funcionamento entre sistemas;
- Software é algo complexo para construir;
- Cada funcionalidade requer um esforço grande para desenvolvimento;
- Evita-se 'reinventar a roda': alguém cria rodas com tamanhos padronizados, e a questão se resume a 'qual roda devo usar';
- Evita-se o retrabalho na implantação de software;
- Possibilita o compartilhamento de recursos e de esforços.

Assim, podemos defini-la como: A Integração funcional entre diferentes ferramentas de *software* ou *hardware*, provida por meio de padronização de serviços.

## Vejamos o sequinte esquema:

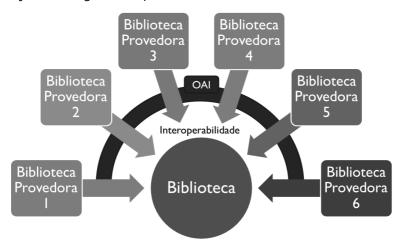


Figura 6 - Interoperabilidade entre Repositórios Digitais

No contexto dos Repositórios Digitais, a interoperabilidade é promovida por tecnologias que regem os processos de comunicação, compartilhamento e troca de dados. Na atualidade a interoperabilidade entre sistemas de Bibliotecas e Repositórios Digitais é pensada em termos de padronização de Metadados e de Protocolos de Comunicação.

#### Padrões de metadados:

Os padrões de metadados são necessários para:

- Facilitar a implantação de sistemas que precisam trocar informações;
- Padronizar o processo de Descrição dos Documentos;
- Permitir que qualquer desenvolvedor crie seu software com padrões de aceitação de mercado.

Por que o uso de 'formatos proprietários' ou particulares é uma má ideia?

- Evitam o compartilhamento de dados;
- Requerem processos de conversão complexos e caros;
- São contrários às boas políticas de Preservação Digital.

Uma das preocupações da área de Ciência da Informação é facilitar o acesso à informação como forma de recuperação e disseminação da informação, e é nesse sentido que o uso de tecnologias existentes são periodicamente aplicadas, no intuito de aprimorar seus benefícios para a área.

Os estudos relacionados aos processos de automação proporcionaram uma melhor interação entre bibliotecas nacionais e internacionais, concretizando assimo intercâmbio informacional.

Um sistema de intercâmbio é capaz de suprir deficiências e aperfeiçoar a aplicação de recursos, e é com base nesse pressuposto que se verifica o compartilhamento de informação como algo inovador, sobretudo para se estabelecer padrões ao enorme volume de documentos eletrônicos vigentes.

#### Conceito de metadados

Os metadados constituem importantes ferramentas para a descrição do conteúdo de um determinado conjunto de dados de um item informacional. A padronização possibilita o fácil acesso e recuperação da informação e os usuários podem mover com facilidade os dados entre os diferentes sistemas e plataformas informáticas.

Mesmo estando atual somente a partir da década de 90, o termo *Metadata*, não é tão recente. A palavra *Metadata* foi criada por Jack Myres em 1969, para denominar os dados que descreviam registros de arquivos convencionais.

Para Alvarenga (2003, p.19), o conceito de metadado,

etimologicamente, quer dizer 'dado sobre dado'; dado que descreve, a essência, atributos e contexto de emergência de um recurso (documento, fonte, etc.) e caracteriza suas relações, visando-se ao acesso e ao uso potencial. O prefixo grego meta significa mudança, posterioridade, além, transcendência [...].

Duval et al. (2002, p.11) definem metadados como sendo dados estruturados sobre dados. Metadados ou Metadata é a descrição de dados sobre dados. De acordo com Garcia (1999, p.3),

podem ser destacadas as seguintes vantagens na utilização e disponibilização de metadados: a) estabelecimento de padrões de dados diante da heterogeneidade de informações contidas na rede, como por exemplo a Internet; b) facilidade e maior precisão na recuperação das informações desejadas, troca de informações entre aplicações e organizações.

Já Souza et al. (2000) apontam como vantagens os sequintes itens:

a) possibilitar a interoperabilidade entre as diversas fontes de dados; b) definir a linguagem de consulta; c) permitir a agilidade e o acesso com qualidade na recuperação da informação; d) propiciar o intercâmbio informacional.

Em síntese, podemos caracterizar o papel dos metadados com um recurso informacional, produto da aplicação de um conjunto de técnicas de representação temática e descritiva, que agrega valor ao documento (no seu sentido mais amplo) e que promove a acessibilidade, a compatibilidade e a capacidade de compartilhamento.

# O padrão Dublin Core

Em 1995 na cidade de Dublin, Ohio, EUA, foi realizado *o I Dublin Metadata Workshop*. O principal resultado desse Workshop foi a criação de um recurso simples de descrição de conteúdos, com potencial para otimizar o acesso à informação na internet.

O Dublin Core Metadata Initiative (DCMI) é um projeto destinado a organizar as informações nas páginas da Web, com o objetivo de estabelecer padrões de catalogação e classificação das informações no meio eletrônico. O DC, tem suas origens em Chicago, na 2a. Conferência Internacional sobre a WWW em outubro de 1994, quando Yuri Rubinsky, Stuart Weibel e Eric Miller integrantes da OCLC – Online Computer Library Center e Joe Hardin da NCSA – National Center for Supercomputing Applications, iniciaram uma discussão em semântica e Web.

Essa iniciativa levou a NCSA e a OCLC a organizarem em 1995 um evento denominado *OCLC/NCSA Metadata Workshop*, onde os participantes discutiram um conjunto semântico para recursos baseados na WEB, com o propósito de agilizar a pesquisa e recuperação de recursos informacionais na Web. O objetivo principal desse

workshop era definir um conjunto mínimo de elementos de descrição para recursos da Web.

O primeiro *workshop*, realizado em Dublin, Ohio, teve três objetivos principais:

- Identificar a clientela ou comunidades envolvidas com recuperação de documentos;
- Imaginar as descrições de metadados em potencial que pudessem atender essa população de usuários, e;
- Chegar a um consenso sobre quais são os elementos descritivos mínimos necessários para facilitar a recuperação da informação.

O DCMI apresenta um conjunto de 15 elementos de metadados (descritos mais adiante), considerados mínimos para facilitar a recuperação da informação do documento eletrônico. Segundo Souza et al. (2000):

O Dublin Core, pode ser definido como sendo o conjunto de elementos de metadados planejado para facilitar a descrição de recursos eletrônicos. A expectativa é de que os autores e Websites, que não possuam conhecimentos em catalogação, possuam capacidade de usar o Dublin Core para descrição de recursos eletrônicos, tornando suas produções mais visíveis aos mecanismos de busca e sistemas de recuperação.

O conjunto de descritores do DC pode estar intrínseco no próprio documento descrito – por meio da linguagem HTML (*Hiper Text Markup Language*), XML e outras, ou, de acordo com o recurso necessário, a meta-informação pode estar separada do recurso utilizado para a catalogação.

Duval  $\it{et~al.}$  (2002) explicam que os metadados DC podem ser usados nas seguintes formas:

 embedded metadata (metadados embarcados no documento) pode ser minerado (ou seja, pode ser rastreado na internet por um motor de busca), o que aumenta sua visibilidade e pode incentivar os criadores a implementar metadados em seu recurso;

Um exemplo de código html com os metadados embarcados no documento é mostrado a seguir:

```
k rel="schema.DC" href="http://purl.org/DC/elements/1.1/">
<META NAME="DC.title" CONTENT="Dublin Core Metadata Initiative (DCMI)
Home Page">
META NAME="DC.description" CONTENT="The Dublin Core Metadata
Initiative is an open forum engaged in the development of interoperable
online metadata standards that support a broad range of purposes and
business models. DCMI's activities include consensus-driven working
groups, global conferences and workshops, standards liaison, and
educational efforts to promote widespread acceptance of metadata
standards and practices.">
<META NAME="DC.date" CONTENT="2004-06-14">
<META NAME="DC.format" CONTENT="text/html">
<META NAME="DC.format" CONTENT="text/html">
<META NAME="DC.contributor" CONTENT="Dublin Core Metadata Initiative">
<META NAME="DC.language" CONTENT="en">
<META NAME="DC.language" CONTENT="text/html; charset=iso-8859-1">
```

Figura 7 - Exemplo de tags DCMI embarcadas no código HTML de um documento da Web

 associated metadata é mantido em arquivos associados ao recurso que descrevem, eles podem ou não serem minerados; a principal vantagem desse tipo de metadado é que ele facilita o controle sem alterar o conteúdo do recurso, mas esse benefício é pago ao custo da simplicidade, exigindo co-gerenciamento dos arquivos de recursos e dos arquivos de metadados.

# Terminologia usada pelo DCMI, segundo a norma ANSI/NISO Z39.85

- DCMI: o Dublin Core Metadata Initiative, mantenedor do Dublin Core Metadata Element Set (conjunto de elementos de metadados Dublin Core)
- Ciclo de vida de um Recurso: uma sequencia de eventos que marcam o desenvolvimento e uso de um recurso. Alguns exemplos

de ciclo de vida são: concepção de uma invenção, criação de um rascunho, revisão de um artigo, publicação de um livro, entre outros.

• Recurso: qualquer coisa que possa ser definida e identificada

Elemento	Nome de Exibição	Definição	Comentário
title	Título	Um nome dado ao recurso	Tipicamente, um Título será o nome pelo qual o recurso é formalmente conhecido.
creator	Criador	A entidade primariamente responsável pela criação do recurso	Exemplos de Criador incluem uma pessoa, uma organização, ou um serviço. Tipicamente, o nome de um Criador deve ser usado para indicar uma entidade.
subject	Assunto	O tema do recurso	Tipicamente, um Assunto deverá ser expresso por palavras chave, frases, ou códigos de classificação que descrevem o conteúdo do recurso. Como boa prática recomenda-se a seleção de termos de vocabulários controlados, ou de sistemas de classificação formais. Para descrever características espaciais ou temporais, use o elemento <i>coverage</i>
description	Descrição	Uma descrição do conteúdo do recurso	Descrições podem incluir, sem estarem limitadas a tal: um resumo, um índice, uma referência a uma representação gráfica do conteúdo, ou uma descrição textual.
publisher	Editor	Uma entidade responsável por tornar o recurso acessível	Exemplos de um Editor incluem uma pessoa, uma organização ou um serviço. Tipicamente, o nome de um Editor deve ser usado para indicar a entidade.
contributor	Contribuidor ou Colaborador	Uma entidade responsável por qualquer contribuição para o conteúdo do recurso.	Exemplos de Contribuinte ou Colaborador incluem uma pessoa, organização ou serviço. Tipicamente, o nome deve ser usado para indicar a entidade.
date	Data	Uma data associada a um evento do ciclo de vida do recurso.	Tipicamente, uma Data deve ser associada à criação ou disponibilidade do recurso. Como boa prática recomenda-se para codificação de valores de datas um perfil da norma ISO 8601 [W3CDTF], segundo o formato AAAA-MM-DD.

type	Tipo do Recurso	A natureza ou gênero do recurso	Tipos incluem termos descrevendo categorias genéricas, funções, gêneros, ou níveis de agregação para o conteúdo. Recomenda-se como boa prática a seleção de valores a partir de vocabulários controlados (por exemplo, a lista do documento de trabalho "Dublin Core Types" [DCT1]). Para descrever a manifestação física ou digital do recurso, deve ser usado o elemento <i>format</i> .
format	Formato	O formato do arquivo, tipo de mídia física ou dimensões do recurso	Tipicamente, o Formato deve incluir o tipo de meio do recurso, ou as suas dimensões. Este elemento deve ser usado para determinar as aplicações informáticas ou qualquer tipo de equipamento necessário para reproduzir ou operar com o recurso. Exemplos de dimensões incluem tamanho e duração. Como boa prática recomenda-se a seleção de valores a partir de vocabulários controlados (como, por exemplo, a lista de "Internet Media Types" [MIME] definindo formatos e meios).
identifier	Identificador do Recurso	Uma referência não ambígua ao recurso, definida num determinado contexto.	Como boa prática recomenda-se a identificação do recurso por meio de uma cadeia de caracteres ou por um número de acordo com um sistema de identificação formal. Exemplos de sistemas de identificação formais incluem o "Uniform Resource Identifier" (URI) (incluindo o "Uniform Resource Locator" (URL)), o "Digital Object Identifier" (DOI) e o "International Standard Book Number" (ISBN), ou identificadores permanentes do Handle System.
source	Fonte	Uma referência a um	O presente recurso pode ser derivado do recurso Fonte na sua

		recurso possa ter derivado.	se a referência ao recurso fonte através de um identificador em conformidade com um sistema de identificação formal.
language	Idioma	O idioma do recurso.	Como boa prática recomenda-se o uso de um vocabulário controlado como a norma RFC 4646, ou ainda, Como boa prática recomenda-se para valores do elemento Língua a utilização do RFC 1766, o qual inclui um código de língua de duas letras (retirado da norma ISO 639), seguido opcionalmente por um código de duas letras para o país (retirado da norma ISO 3166). Por exemplo, 'en' para Inglês, 'fr' Francês, ou 'en-uk' para o Inglês do Reino Unido.
relation	Relação	Uma referência a um recurso relacionado.	Como boa prática recomenda-se referir o recurso através de uma cadeia de caracteres ou número em conformidade com um sistema de identificação formal.
coverage	Cobertura	O escopo espacial ou temporal do recurso, a aplicabilidade espacial do recurso ou a jurisdição sob a qual o recurso está sujeito.	Cobertura inclui tipicamente uma localização espacial (o nome de um lugar ou coordenadas geográficas), um período no tempo (a sua designação, data, ou intervalo de tempo), ou jurisdição (o nome de uma entidade administrativa). Como boa prática recomenda-se a seleção de valores de vocabulários controlados (como por exemplo, o "Thesaurus of Geographic Names"), devendo ainda ser usados, quando apropriado, preferencialmente nomes de lugares e designações de períodos no tempo, em vez de identificadores numéricos tais como coordenadas ou intervalos de datas.
rights	Direitos Autorais	Informação sobre a propriedade intelectual sobre o recurso	Tipicamente inclui informações sobre propriedade intelectual, direitos de autor e afins.

#### Qualificadores do Dublin Core

Os elementos denominados *qualificadores* especializam a semântica dos elementos básicos do DC, detalhando a descrição do conteúdo que por vezes pode ser muito genérica. A especificação completa está disponível em http://dublincore.org/documents/2000/07/11/dcmesqualifiers/. Vejamos no quadro a seguir a especificação dos qualificadores:

DCMES Element	Element		Element
	Refinement(s)		Encoding Scheme(s)
Title	Alternative		-
Creator	-		-
Subject	-		LCSH MeSH (DeCS) DDC LCC UDC
Description	Table Of Conte Abstract	nts	-
Publisher	-		-
Contributor	-		-
Date	Created Valid Available Issued Modified		DCMI Period W <sub>3</sub> C-DTF
Type	-		DCMI Type Vocabulary
Format	Extent		
	Medium		IMT
Identifier	-		URI
Source	-		URI
Language	-		ISO 639-2 RFC 1766
Relation	Is Version Of Has Version Is Replaced	Is Part Of Has Part Is	URI

	By Replaces Is Required By Requires	Referenced By References Is Format Of Has Format	
Coverage	Spatial		DCMI Point ISO 3166 DCMI Box TGN
	Temporal		DCMI Period W <sub>3</sub> C-DTF
Rights			

#### O Protocolo OAI-PMH

A origem do Protocolo OAI-PMH (*Open Archives Initiative Protocol for Metadata Harvesting*) é parte da Iniciativa *Open Archives* (Arquivos Abertos), ou OAI, originada da Convenção de Santa Fé (Novo México) no final de 1999, objetivando promover soluções de interoperabilidade que facilitem uma disseminação eficiente dos conteúdos digitais, com caraterísticas técnicas livres de restrições econômicas e independentes de plataforma.

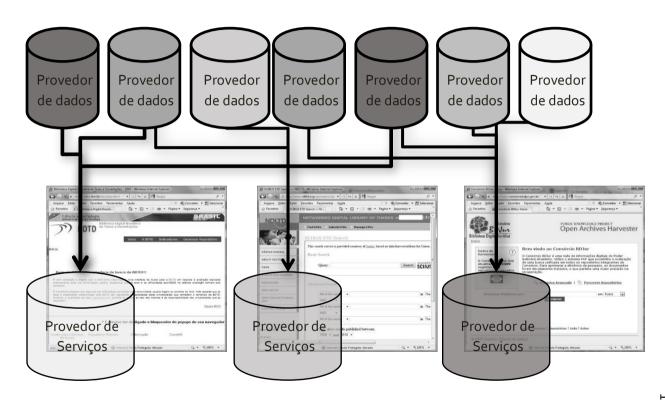
Trata-se de um padrão aberto, do ponto de vista de sua arquitetura, não significando acesso gratuito ou ilimitado às informações constantes dos repositórios que fazem parte da Iniciativa, em outras palavras, o OAI-PMH é que é aberto.

A base da Iniciativa é a possibilidade de compartilhar seus metadados, por meio de processos que regem o intercâmbio (protocolo), com papéis e tecnologias previamente definidos.

Os participantes da Iniciativa são divididos em Provedores de Dados (*Data Providers* ou DP) e Provedores de Serviços (*Service Providers* ou

SP). Os provedores de dados mantêm repositórios de documentos digitais que implementam o protocolo OAI-PMH como forma de expor os metadados de seus documentos. Já os provedores de serviços oferecem buscas a estes metadados ou outros serviços que visam agregar valor à Iniciativa, além de funcionarem como provedor de dados.

A figura a seguir ilustra essa arquitetura:



O protocolo OAI-PMH introduz o conceito de *Metadata Harvesting* (em português: colheita de metadados), um processo no qual os provedores de serviços, a partir da lista de repositórios dos provedores de dados, realizam periodicamente uma busca a estes provedores de dados, "colhendo" e armazenando os metadados para exibição sob a forma de consultas efetuadas pelos usuários.

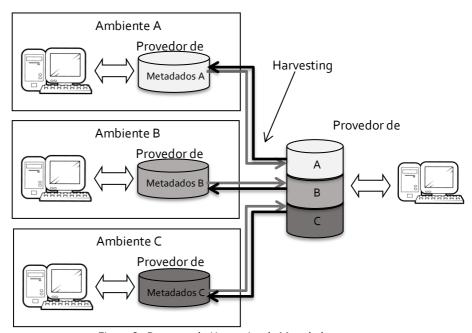


Figura 8 - Processo de Harvesting de Metadados

Dessa maneira torna-se possível agregar e conectar coleções de diferentes documentos, de diferentes locais e de diferentes plataformas.

A comunicação entre os hosts é regida por um protocolo de requisições e respostas definidas na forma de comandos (denominados verbos<sup>7</sup>) que são transmitidos via URI no endereço do Repositório. Vejamos um exemplo:

Abra um navegador de sua preferência, copie e cole a seguinte URL na barra de endereço: http://www.labtecgc.udesc.br:8081/tabdoai/request?verb=Identify

A resposta enviada pelo Provedor de Dados é sempre na forma XML, como se segue:

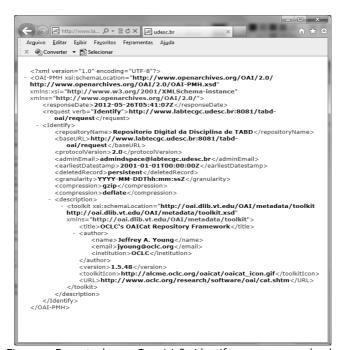


Figura 9 - Reposta de uma Requisição Identify para um provedor de dados

<sup>&</sup>lt;sup>7</sup> A especificação completa do OAI-PMH está disponível em: http://www.openarchives.org/OAI/openarchivesprotocol.html

O intercâmbio de metadados dos documentos é realizado na forma de um "diálogo" entre os hosts envolvidos, conferindo a confiabilidade e integridade necessárias.

Esse processo precisa ser configurado para que ele ocorra de maneira automática. Por exemplo, pode-se escolher somente determinadas coleções e a política de atualização, que pode ser somente para itens modificados (com ou sem exclusão) ou para a coleção inteira.

A seguir, para fins de ilustração, um quadro com a lista de verbos que os sistemas que implementam o OAI-PMH utilizam para realizar a interoperabilidade:

Verbo	Descrição	Argumentos
GetRecord	Recupera os metadados de um item individual de um repositório.	identifier. Obrigatório. Com ele, definimos o identificador único do item de um repositório. metadataPrefix. Obrigatório. Define o padrão de metadados adotado que deve estar especificado no Provedor de Dados.
Identify	É usado para coletar informações sobre um repositório.	Não há argumentos.
ListRecords	Este verbo recupera os metadados de um repositório.	from. Opcional. Os dados coletados devem ser criados ou alterados a partir da data específica por este argumento.  until. Opcional. Os dados coletados devem ser criados ou alterados até a data especificada pelo argumento.  metadataPrefix. (o mesmo listado em GetRecord).  set. Opcional. Especifica um conjunto, para o havester poder refinar a sua coleta.  resumptionToken. Exclusivo. Argumento necessário quando os provedores utilizam o controle de fluxo na coleta dos metadados.
ListIdentifi ers	É uma abreviação do <b>ListRecords,</b> que retorna apenas	from. until. metadataPrefix. set. resumptionToken.

	o <i>header</i> de um repositório.	
ListMetada	Retorna os padrões	identifier. Opcional (apenas neste verbo).
taFormats	de metadados utilizados em um	Retorna o padrão de metadados utilizado em um item específico.
	repositório.	
ListSets	É utilizado para	resumptionToken.
	retornar a estrutura	
	de um repositório,	
	listando todos os	
	conjuntos que	
	compõe os	
	metadados	

Quadro 1 - Verbos do OAI-PMH e seus argumentos - Cardoso Jr. (2007)

# Considerações finais

Repositórios e Bibliotecas Digitais podem ser vistos como 'paradigmas' em evolução. A preservação do conhecimento, do patrimônio cultural, científico e histórico está fortemente associada a investimentos em infraestruturas e serviços baseados nessa forma de preservação, memória e disseminação da informação e conhecimento.

Muitas instituições ainda têm dificuldades em absorver processos de gestão e conhecimento tecnológico para manter serviços de acesso à informação e documentos digitais por meio de repositórios e bibliotecas digitais.

A realidade dessas iniciativas é caracterizada por empreitadas de natureza multidisciplinar, envolvendo profissionais da Ciência da Informação, Tecnologias da Informação e Administração. Somente a conjunção de habilidades e competências dessas três áreas viabilizam a criação e a sustentabilidade de um projeto que objetiva, em primeira instância, apoiar a digitalização, acesso e preservação do patrimônio cultural e científico, respeitando as questões relativas à propriedade intelectual, com sistemas computacionais capazes de 'sobreviver' à obsolescência tecnológica.

#### REFERÊNCIAS

ALVARENGA, LÍDIA. Representação do Conhecimento na Perspectiva da Ciência da Informação em Tempo e Espaço Digitais. *Encontros Bibli*: Revista Eletrônica de Biblioteconomia e Ciência da Informação, Florianópolis, v. 8, n. 15, p.18-40, 2003. Disponível em:<a href="http://www.periodicos.ufsc.br/index.php/eb/article/view/97/5233">http://www.periodicos.ufsc.br/index.php/eb/article/view/97/5233</a>. Acesso em: 01-ago-2008.

BOSO, AUGIZA KARLA. Repositórios de instituições federais de ensino superior e suas políticas: análise sob o aspecto das fontes informacionais. (2011). 150 f. dissertação (Mestre em Ciência da Informação) - Departamento de Ciência da Informação, Universidade Federal de Santa Catarina, Florianópolis, 2011.

CAFÉ, LÍGIA et al. Repositórios institucionais: nova estratégia para publicação científica na Rede. In: XXVI Congresso Brasileiro de Ciências da Comunicação, 02 a o6 de setembro de 2003, Belo Horizonte - MG 2003, Conference Proceedings. Place. Published: PUC-MG, 2003. p.1-12. Disponível em:<a href="http://dspace.ibict.br/dmdocuments/ENDOCOM\_CAFE.pdf">http://dspace.ibict.br/dmdocuments/ENDOCOM\_CAFE.pdf</a>>. Acesso em: 01-ago-2011.

CARDOSO JR., MARCOS JOSÉ DE M. *Clio-i*: interoperabilidade entre repositórios digitais utilizando o protocolo OAI-PMH. (2007). 137 f. Dissertação (Mestre em Ciência da Computação) - Centro de Informática, Universidade Federal de Pernambuco, Recife, 2007. Disponível em:<a href="http://www.cin.ufpe.br/~rbcp/dissertacoes/dissertacao">http://www.cin.ufpe.br/~rbcp/dissertacoes/dissertacao</a> CARDOSO.pdf>. COSTA, SELY MARIA DE SOUZA; LEITE, FERNANDO CÉSAR LIMA. Insumos conceituais e práticos para iniciativas de repositórios institucionais de acesso aberto à informação científica em bibliotecas de pesquisa. *Implantação e gestão de repositórios institucionais*. Salvador - BA: EDUFBA, 2009. p. 163-202. Disponível em:<a href="http://www.edufba.ufba.br/2010/09/implantacao-e-gestao-de-repositorios-institucionais/">http://www.edufba.ufba.br/2010/09/implantacao-e-gestao-de-repositorios-institucionais/>.

DRABENSTOTT, KAREN M.; BURMAN, CELESTE M. Revisão analítica da biblioteca do futuro. *Ciência da Informação*, v. 26, 1997. Disponível em:<a href="http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200012&nrm=iso>">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S010000000012&nrm=iso>">http://www.scielo.br/scielo.php.nrm=iso>"

DUVAL, ERIK et al. Metadata Principles and Practicalities. *D-Lib Magazine*, v. 8, n. 4, 2002. Disponível em:<a href="http://www.dlib.org/dlib/aprilo2/weibel/04">http://www.dlib.org/dlib/aprilo2/weibel/04</a>

weibel.html>. Acesso em: 01/03/2012.

GARCIA, SIMONE DE SOUZA. *Metadados para documentação e recuperação de imagens*. (1999). 152 f. dissertação (Mestrado), Instituto Militar de Engenharia, 1999. Disponível em:<a href="http://www.ipanema.ime.eb.br/dissertacoes/1999-Simone\_Garcia.pdf">http://www.ipanema.ime.eb.br/dissertacoes/1999-Simone\_Garcia.pdf</a>>.

KURAMOTO, HÉLIO. Informação científica: proposta de um novo modelo para o Brasil. *Ciência da Informação*, v. 35, p. 91-102, 2006.

MIRANDA, ANTONIO. Os Conceitos de Organização Baseada na Informação e no Conhecimento e o desenvolvimento de Serviços Bibliotecários. *Ciência da Informação*, v. 22, n. 3, p. 227-232, 1993.

REZENDE, YARA. Natura cosméticos: quando é virtuoso ser virtual. Ciência da Informação, v. 26, 1997. Disponível em:<a href="http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200007&nrm=iso">http://www.scielo.br/scielo.php?script=sci\_arttext&pid=S0100-19651997000200007&nrm=iso</a>.

RIBEIRO JR., DIVINO IGNACIO RIBEIRO. Agentes inteligentes como mediadores na recuperação da informação. (2001). 113 f. dissertação (Mestrado) - Departamento de Ciência da Informação, Universidade Estadual Paulista 'Júlio de Mesquista Filho' - UNESP, Marília, 2001.

SAYÃO, LUIS et al. *Implantação e Gestão de Repositórios Institucionais*: políticas, memória, livre acesso e preservação. EDUFBA, 2009. 365 p.

SOUZA, MARCIA IZABEL FUGISAWA et al. Metadados para a descrição de recursos de informação eletrônica: utilização do padrão Dublin Core. *Ciência da Informação*, v. 29, p. 93-102, 2000.

# Adaptação de conteúdo para Dispositivos Móveis

Fábio Bif Goularte<sup>1</sup>; Beatriz Wilges<sup>2</sup>; Silvia Modesto Nassar<sup>3</sup>

# 1- Introdução

A transformação do paradigma da *Web* 1.0 para a *Web* 2.0 incide sobre os dispositivos móveis, pois as novas tecnologias de informação e comunicação estão gerando uma mudança na arquitetura de apresentação para uma arquitetura de participação. Os modelos lineares e fechados, presentes na *Web* 1.0, migram para modelos abertos e dinâmicos, configurando um novo cenário, com possibilidades significativas de aprendizagem para estudantes (SINCLAIR et al., 2006). Assim, a construção de conteúdos para as tecnologias móveis deve considerar o contexto da Web 2.0 (GOULARTE, WILGES, NASSAR, 2013).

<sup>&</sup>lt;sup>1</sup> Mestrando do Programa de Pós Graduação em Ciência da Informação da Universidade Federal de Catarina (PPGCIN/UFSC). CONTATO: fabio.bif@posgrad.ufsc.br

<sup>&</sup>lt;sup>2</sup> Professora do Curso Superior de Tecnologia de Redes de Computadores do Serviço Nacional de Aprendizagem Industrial de Santa Catarina (SENAI-SC), doutoranda do Programa de Pós-Graduação em Engenharia e Gestão do Conhecimento da Universidade Federal de Catarina (PPEGC/UFSC). CONTATO: beaw@inf.ufsc.br

<sup>&</sup>lt;sup>3</sup> Professora do Programa de Pós Graduação em Ciência da Informação da Universidade Federal de Catarina (PPGCIN/UFSC), coordenadora do Projeto e-TEC de Acompanhamento e Validação de Materiais Didáticos (UFSC). CONTATO: silvia@inf.ufsc.br

Segundo Tim (2005), o termo *Web* 2.0 se refere a uma mudança da internet para plataforma, onde os aplicativos ou serviços desenvolvidos devem aproveitar os efeitos de rede, tornando-se melhores à medida que forem utilizados pelas pessoas.

No contexto evidenciado, a leitura ou aprendizado de conteúdo por dispositivos móveis requerem o conceito de aprendizagem móvel (*m-learning*). Segundo Saccol, Schlemmer e Barbosa (2011) a aprendizagem móvel se refere a processos de aprendizagem apoiada pelo uso de tecnologias de informação e comunicação, onde a principal característica é a mobilidade. Através de um dispositivo móvel, uma pessoa pode fazer uso de inúmeros recursos no momento e local mais oportuno.

Na aprendizagem móvel, a mobilidade estabelece uma facilidade de acesso à informação. A informação é acessível em qualquer tempo e espaço, e isso pode permitir uma maior autonomia no processo de aprendizagem. Por exemplo, um estudante, além de acessar ou capturar dados onde quer que ele se encontre, também pode transformá-los em informação quase instantaneamente. A informação torna-se mais "presente" e assume um papel de destaque, juntamente com a mobilidade.

Scanlon et al. (2005) descrevem três fatores para delinear a maneira pela qual a aprendizagem móvel é concebida: os estudantes estão em movimento, tanto fisicamente como de outras formas, uma grande quantidade de aprendizagem ocorre fora da situação de aprendizagem formal e a natureza ubíqua da aprendizagem.

Contudo, existem alguns desafios de implementação inerentes à própria tecnologia, e questões como capacidade de processamento, tamanho da tela e adaptação de aplicações desenvolvidas para a *Web* que precisam de atenção (MANDULA et al., 2013; ALLY, 2009). Além disso, para proporcionar uma interface de interação humanocomputador (IHC) apropriada para dispositivos móveis, é necessário

considerar padrões que atendam as expectativas dos processos de aprendizagem.

A fim de explorar novas tecnologias para promover o interesse na leitura e aprendizagem em meio digital, este artigo descreve formas de adaptação de conteúdo para dispositivos móveis. Apresenta-se ainda um estudo de caso que resultou em um protótipo de uma ferramenta desenvolvida sob essa nova abordagem para experimentação e validação. Dessa forma, este trabalho pretende contribuir com as áreas de pesquisas relacionadas a aprendizagem em dispositivos móveis.

Assim, este artigo está organizado da seguinte forma: a segunda seção descreve sobre trabalhos relacionados. A terceira seção descreve características, abordagens observadas no desenvolvimento de materiais para dispositivos móveis, bem como tecnologias. Na quarta seção é apresentado um protótipo de material didático adaptado para dispositivos móveis e, por fim, são apresentadas as considerações finais.

#### 2- Trabalhos relacionados

Na pesquisa bibliográfica realizada não foi encontrado nenhum trabalho com o mesmo foco do modelo proposto neste artigo. Por esse motivo, foram explorados os trabalhos que propunham soluções que favorecem a interatividade com o conteúdo e facilitassem a leitura digital.

O trabalho de Mühlbeier et al. (2012) aborda o tema sobre o Novo Acordo Ortográfico, para países de língua portuguesa, em vigor desde Janeiro de 2009. O aplicativo proposto apresenta o tema na forma de histórias em quadrinhos onde o conteúdo é acessado, quadro a quadro, por botões de navegação.

Duarte et al. (2013) demonstram em sua pesquisa o desenvolvimento de um protótipo de aprendizagem móvel à saúde, com foco no ensino e treinamento sobre endometriose. O protótipo foi avaliado com a ajuda de especialistas (médicos) e usuários comuns, sendo identificados tanto aspectos positivos como negativos em sua utilização.

A pesquisa de Morais e Jaques (2013) tem por objetivo desenvolver uma ferramenta para auxiliar na conversão de unidades e na resolução de equações algébricas. A ferramenta funciona como um Sistema Tutor Inteligente e que utiliza técnicas de Inteligência Computacional para promover o ensino de estudantes.

# 3- Abordagens investigadas

Inicialmente, duas possibilidades para adequar conteúdo a dispositivos móveis foram investigadas, uma com a abordagem de livros digitais, os *ebooks*, e outra com a abordagem de aplicativos.

# 3.1. Aplicações baseadas no padrão ePub

O formato de arquivo *ePub* (Electronic Publication) é o formato oficial para distribuição e venda de livros digitais. O *ePub* é um padrão internacional, livre e aberto, o que permite o seu aperfeiçoamento à medida que o mercado e demais tecnologias evoluem.

Um arquivo *ePub* pode ser lido por vários dispositivos que existem hoje no mercado, além disso, o texto é redimensionado automaticamente para o tamanho da tela, melhorando a experiência de leitura.

Atualmente o *ePub* está na versão 3, onde muitas melhorias foram adicionadas, comparadas as versões anteriores, especialmente relacionadas à interatividade. Devido à adoção do HTML5, houve melhora significativa quanto ao formato de publicação e suporte a conteúdos, tais como: vídeo, áudio, *scripting*, imagens, fontes, interatividade digital entre outros.

Contudo, ressalve-se que embora o ePub3<sup>4</sup> esteja evoluindo para um padrão mundial, ainda é uma tecnologia que necessita de mais tempo e pesquisa para ser suportada pela grande variedade de dispositivos e sistemas móveis presentes no mercado.

# 3.2. Aplicações móveis: ferramentas e tecnologias

Existem aplicações móveis que são nativas e as baseadas na Web. Aplicações nativas são mais sofisticadas, pois permitem controlar e utilizar de forma plena os recursos que o dispositivo móvel dispõe (câmera, vídeo, áudio, GPS), mas requerem maiores custos (JUNTUNEN, JALONEN e LUUKKAINEN, 2013).

O espaço mobile é fragmentado, pois as aplicações estão atreladas a sistemas específicos, tais como: Android, iOS, Bada, BlackBerry, Symbian, webOS e Windows Phone. Sendo assim, para atingir uma maior base de usuários, é necessário desenvolver aplicativos para diferentes sistemas operacionais e, para tal, é necessária uma equipe de profissionais altamente qualificada, o que significa maior tempo e recursos de desenvolvimento. A fragmentação também existe dentro de um único sistema operacional, considerando que existem novas versões que são lançadas no mercado o que pode ocasionar problemas no funcionamento de aplicativos antigos.

Para contornar algumas das dificuldades mencionadas, há a abordagem de aplicações móveis baseadas na *Web*. Aplicativos baseados na Web são acessados por um browser padrão, mas se comportam como aplicações nativas. Dessa forma, dispensam a necessidade de adaptar a aplicação para cada um dos sistemas operacionais existentes (JUNTUNEN; JALONEN E LUUKKAINEN, 2013). Além disso, a

<sup>&</sup>lt;sup>4</sup> http://www.infodocket.com/2013/03/12/international-publishers-association-endorses-epub-3-as-global-publishing-standard/

computação móvel em nuvem<sup>5</sup> pode ajudar a resolver as limitações de hardware dos dispositivos móveis, executando uma parte do processamento na nuvem (ABOLFAZLI et al., 2013).

O padrão HTLM5 é o responsável por reforçar as aplicações móveis baseadas na *Web*.

HTMI

O HTML é uma linguagem para estruturação e apresentação de conteúdo para a Web. A versão 5, encontra-se em fase de projeto e está sendo desenvolvido pelo Consórcio W3C<sup>6</sup> e pelo WHATWG<sup>7</sup>. Mesmo em fase de projeto, a quantidade de dispositivos móveis com suporte a HTML5 tende a aumentar significativamente para os próximos anos e, com isso, muitas empresas e setores de serviços serão beneficiados.

Em 2016, mais de 2,1 bilhões de celulares terão browsers com HTML5, contra apenas 109 milhões em 2010, 336 milhões em 2011 e cerca de 1 bilhão em 2013 (SHAH, 2011; ABIResearch, 2011).

<sup>&</sup>lt;sup>5</sup> Mobile Cloud Computing (MCC) é um paradigma da computação distribuída composto pela computação móvel, computação em nuvem e redes sem fio com o objetivo de aumentar a capacidade computacional de dispositivos móveis com recursos limitados [weekpedia]

<sup>6</sup> http://www.w3.org/TR/html5/

<sup>&</sup>lt;sup>7</sup> www.whatwg.org

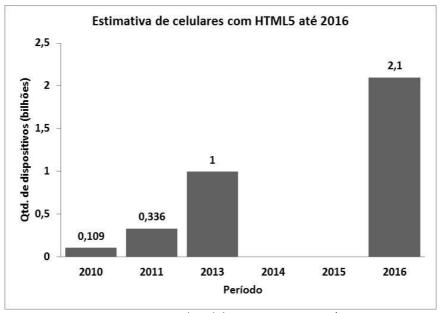


Figura 1. Estimativa de celulares com HTML5 até 2016

Conforme a Figura 1 é possível observar a estimativa de aparelhos celulares com HTML5, no mundo, de 2010 a 2016. Deixando evidente a tendência de aumento, ao longo do tempo, à medida que se transforma em uma linguagem de padrão universal.

Existem várias tecnologias que estão relacionadas ao HTML5, mas as principais são: CSS3 e JavaScript. Uma definição grosseira seria: o HTML5 é usado para manipular o conteúdo, CSS3 para a apresentação do conteúdo e JavaScript para definir o comportamento das outras duas.

# Cascading Style Sheets (CSS)

Cascading Style Sheets (CSS) é um mecanismo simples para adicionar estilo, por exemplo, fontes, cores, espaçamento a um documento web (CSS, 2013). O CSS permite a separação do conteúdo de apresentação

do conteúdo do documento. A última versão, CSS<sub>3</sub>, é totalmente compatível com versões anteriores. Além disso, CSS<sub>3</sub> é dividido em módulos, e mais de cinquenta módulos foram publicados pelo Grupo de Trabalho CSS<sup>8</sup>.

#### JavaScript (JS)

JavaScript é uma linguagem de programação para a web baseada em script. Surgiu da linguagem ECMAScript, padronizada e desenvolvida pela ECMA<sup>9</sup>. Um script é o código do programa que não precisa de préprocessamento, pois é executado pelo browser. Isso torna a página Web mais dinâmica, pois permite manipular o conteúdo sem recarregar uma nova versão de uma página. Esta interatividade adicional faz com que páginas da Web se comportem como um aplicativo de software tradicional (W3C, 2013).

# PhoneGap

Para se criar aplicações móveis multiplataformas, há *frameworks* de desenvolvimento que unem tecnologias comuns para promover uma funcionalidade genérica (RAMADATH e COLLINS, 2012). Algumas são ferramentas proprietárias, sendo necessária uma licença, enquanto outras são livres e de código fonte aberto.

PhoneGap<sup>10</sup> é uma ferramenta de desenvolvimento de código fonte aberto, o que significa que empresas e programadores podem usá-lo na construção de aplicações livres ou comercias para os principais sistemas móveis. Além disso, é uma tecnologia de contêiner que possibilita escrever aplicativos nativamente instalados em dispositivos móveis, utilizando HTML5, CSS3 e JS (ALLEN, GRAUPERA e LUNDRIGAN, 2010) (PALMIERI, INDERJEET e CICCHETTI, 2012). O *framework* JS adotado para o projeto foi o jQuery Mobile.

-

<sup>&</sup>lt;sup>8</sup> http://www.w3.org/Style/CSS/members

<sup>&</sup>lt;sup>9</sup> http://www.ecma-international.org/

<sup>10</sup> http://phonegap.com/

# 3.3. Recomendações e boas práticas de desenvolvimento

Outro fator relevante que deve ser considerado na produção de uma aplicação móvel envolve boas práticas e recomendações de desenvolvimento. São recomendações que diminuem as chances de resultados frustrados, tais como, a aplicação final não estar de acordo com as especificações do projeto. Para isso, Woods e Will (2012), consideram o seguinte:

- Siga uma abordagem ágil usando janelas de desenvolvimento curtas;
- Assegure a entrada de design gráfico adequado ao longo do projeto;
- Estabeleça as características e prioridades mais adequadas para o desenvolvimento da aplicação;
- Quando for o caso, assegure que haja uma adequada extensão e APIs para construir o aplicativo;
- Projete para uma única plataforma, inicialmente, garantindo que considere as questões sobre normas e interface específicas.

Mandula et al. (2013) também apresentam orientações para profissionais e pesquisadores que trabalham com foco na educação criando experiências de ensino que tornam a aquisição de conhecimento mais eficiente, eficaz e atraente. As recomendações são:

 Entender o estudante: é preciso conhecer as necessidades dos utilizadores finais da ferramenta e o contexto em que será utilizada;

- Limitar o conteúdo gráfico: muito conteúdo gráfico pode aumentar a carga de processamento e drenar a energia da bateria. Deve-se ter atenção com o conteúdo em formato flash porque não têm suporte em todos os dispositivos móveis;
- Encurtar e fragmentar o conteúdo: tendo em vista as limitações no tamanho da tela e no processamento dos dispositivos móveis, deve-se pensar em estratégias para que o conteúdo atinja um número máximo de estudantes móveis. Por exemplo, estudantes não prestam atenção em vídeos por mais de 10 e 15 minutos por causa da tela menor;
- Suporte multimídia: a maioria dos dispositivos móveis têm diferentes recursos de hardware, portanto, pode ser difícil apoiar todos os formatos de mídia;
- Design interativo: a interface do usuário deve ser simples e envolvente, permitindo a interatividade e compreensão do assunto e envolvendo o estudante no processo de educação;
- Dimensões móvel: o projeto tem que ser compatível com vários dispositivos. Por exemplo, um *smartphone* tem dimensões diferentes quando comparado com *tablets*, e aliado a isso há uma grande variedade de produtos no mercado.

Devido à grande quantidade de referências na literatura e à variedade de tecnologias esta seção descreveu apenas sobre as principais tecnologias e recomendações utilizadas no protótipo proposto, que será apresentado na seção seguinte.

## 4- Um protótipo de material didático para dispositivos móveis

Para reduzir a complexidade e os esforços iniciais, e por se tratar de uma nova abordagem para lidar com a leitura e aprendizagem de conteúdo em meio digital decidiu-se concentrar o desenvolvimento do protótipo em uma única plataforma móvel.

Android foi o sistema escolhido, com vista a explorar as outras plataformas em uma fase posterior. A escolha se deve ao grande número de dispositivos existentes no mercado, a natureza aberta da tecnologia, as pesquisas no meio acadêmico e, principalmente, por ser o sistema adotado nos *tablets* distribuídos pelo governo nas escolas públicas brasileiras.

Com relação ao conteúdo a ser adaptado, foi escolhido um caderno didático, em formato PDF, elaborado para um curso técnico de um Instituto Federal em parceria com a Rede e-Tec Brasil<sup>11</sup>. A Rede e-Tec Brasil é uma ação do Ministério da Educação (MEC), que visa a oferta de educação profissional e tecnológica na modalidade à distância (WILGES et al., 2012). O material escolhido é de uma disciplina que também é comum a outros cursos técnicos e consequentemente a outros eixos tecnológicos.

A implementação do caderno para dispositivos móveis segue a arquitetura proposta na Figura 2. A arquitetura proposta está baseada em três níveis e isso se deve a estruturação do projeto e da equipe e-Tecna UFSC.

\_

<sup>11</sup> http://redeetec.mec.gov.br/

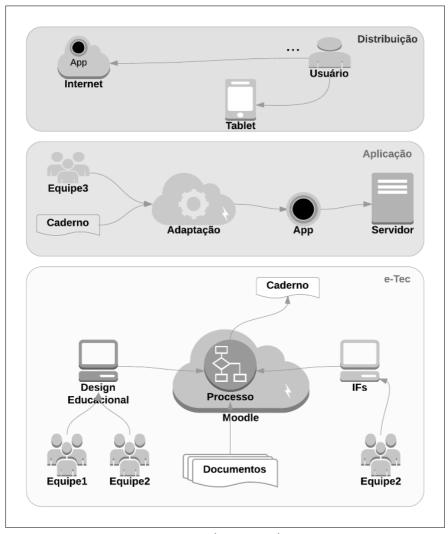


Figura 2. Arquitetura proposta com base na Rede e-Tec em parceria com a equipe UFSC

Conforme Figura 2, na base, encontram-se todos os componentes envolvidos na concepção de um caderno didático. Fazem parte desse nível a instituição de ensino a qual o caderno didático se destina, a

Comissão de Validação de Mateiras Didáticos (CVMD) da UFSC, com pessoas em papéis distintos e as tecnologias envolvidas no processo. No nível intermediário, encontra-se a fase de adaptação de conteúdo para criação da aplicação móvel. No nível superior, encontra-se o aluno que acessa um serviço e instala o caderno no *tablet*.

## 4.1. Modelo proposto

Adequar uma grande quantidade de informações sobre os conteúdos de disciplinas de um curso técnico para o aprendizado móvel representa um grande desafio. A estratégia adotada para essa situação foi a de representar o conteúdo como uma árvore hierárquica conforme Figura 3. A ideia relacionada à árvore é a de fragmentar o conteúdo dos cursos e das disciplinas.

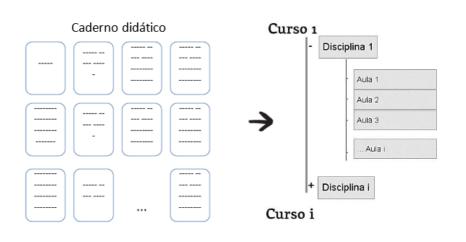


Figura 3. Estratégia de fragmentação

São muitas as disciplinas que compõem um curso, sendo que estruturar todo o conteúdo dessas disciplinas apenas em uma aplicação é inviável e não recomendado. Portanto, cada disciplina representa uma aplicação

móvel e dentro de cada aplicação o conteúdo é dividido em módulos, que são aulas.

O passo seguinte foi criar o design do protótipo através de wireframes. Essa etapa consiste em facilitar o entendimento de como as informações estarão dispostas na aplicação. Normalmente, um wireframe parece como um rascunho, porém contém claramente os principais grupos de conteúdo, a estrutura da informação e uma descrição e visualização básica de uma interface de interação. A Figura 4 apresenta o design das views do protótipo.

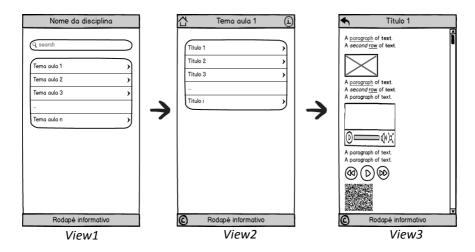


Figura 4. Wireframes do protótipo

Todas as informações do caderno didático estão presentes no protótipo e dispostas em três *views*. Isso ajuda na navegação pelo conteúdo e na interação com o dispositivo móvel, pois se houver muitas *views*, o estudante terá mais esforço para acessar e utilizar o conteúdo no dispositivo, aumentando as chances de uma aprendizagem ineficiente.

Cada view contém três estruturas: cabeçalho, corpo e rodapé. A primeira view apresenta no cabeçalho o nome da disciplina, no corpo

central um campo de pesquisa e os temas de cada aula. A segunda *view* apresenta no cabeçalho o tema da aula e dois botões, e no corpo os títulos do assunto. Na terceira *view* é onde o conteúdo realmente se concentra. No cabeçalho há o título do assunto da aula e um botão para voltar ao tema da aula, e no corpo são utilizadas mídias (texto, imagens, áudio, vídeo) para compor o material. No rodapé de todas as *views* são apresentadas as informações de *copyright*.

## 4.2. Funcionamento da aplicação

O funcionamento do sistema ocorre da seguinte forma: ao acessar a ferramenta, a tela de inicialização apresenta informações, sobre a equipe e instituições responsáveis pela produção do material. Após isso, a navegação pela ferramenta segue o fluxo apresentado na Figura 5.

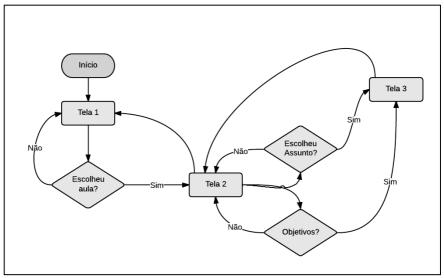


Figura 5. Fluxo de navegação

Na primeira tela o estudante deve escolher qual aula acessar, sendo direcionado para outra tela, onde escolhe o assunto da aula. Ainda na segunda tela, o estudante pode usar dois botões, um para voltar à tela

anterior e outro para apresentar quais são os objetivos da aula. Os objetivos aparecem em uma janela (*lightbox*) com um botão que redireciona para o primeiro assunto da aula, na terceira tela. Na terceira tela, o estudante pode retornar apenas para a tela anterior e não para a inicial, pois deseja-se guiá-lo para que acesse todos os assuntos da aula selecionada antes de acessar uma outra aula. Essa forma de guiar é sutil, pois a ferramenta deve manter uma navegação livre e flexível para o estudante ter liberdade de escolha.

# 4.3. Avaliação preliminar

No momento, o projeto do protótipo encontra-se em fase de teste e avaliação. Uma avaliação preliminar foi realizada entrevistando alguns estudantes (total de 7) que utilizaram a ferramenta para teste. A principal vantagem da avaliação preliminar é fornecer um retorno das impressões do usuário sobre o uso da ferramenta, e com isso, pode-se melhorar a qualidade para a próxima interação em uma avaliação abrangente.

O critério de seleção dos estudantes foi que tivessem conhecimento prévio do material impresso que foi selecionado.

O protótipo foi instalado em um *tablet* Positivo Ypy 7, com tela de 7 polegadas e resolução de 1024 x 768 pixels, sistema operacional Android 2.3, processador de 1GHz e memória RAM de 340MB. A escolha desse modelo se deve ao fato de ser esse um dos primeiros *tablets* que foram distribuídos às escolas públicas. Atualmente o referido modelo sofreu alterações de *design* e houve melhorias na configuração do *hardware*.

O *tablet* com o material didático instalado foi entregue para os estudantes testarem e, em outro momento, responderam a perguntas que foram embasadas na norma ISO 9126, que tem como objetivo validar a qualidade de *software* (ABNT, 2003). As perguntas estão descritas na Tabela 1. Os pontos observados foram: usabilidade,

apreensibilidade, inteligibilidade, atratividade, eficiência, adequação e aspecto pedagógico.

Pergunta	Pontos observados	
O protótipo de material didático para dispositivo móvel é difícil de ser usado?	Usabilidade	
2. O protótipo apresenta dificuldade de navegação ou dificuldade na compressão de seus elementos (botões, texto, imagens entre outros)?	Apreensibilidade e inteligibilidade	
3. Os recursos disponíveis no protótipo satisfazem ao propósito de leitura e aprendizado móvel ou podem ser adicionados novos?	Atratividade	
4. A velocidade de execução da ferramenta e o tempo de resposta são aceitáveis?	Eficiência	
5. Os recursos do protótipo cumprem com a sua função?	Adequação	
6. A disposição do conteúdo e os recursos utilizados para abordá-lo auxiliam o estudante a compreender as aulas da disciplina?	Aspecto pedagógico	

Tabela1. Questões avaliadas

Os pontos observados e que chamaram a atenção foram apreensibilidade, inteligibilidade e atratividade.

Com relação à apreensibilidade e inteligibilidade, buscou-se avaliar a capacidade do protótipo de ser compreendido, seu funcionamento aprendido e a facilidade com que o estudante pode compreender as suas funcionalidades. Apenas um estudante apontou dificuldade para navegar pelo conteúdo. Possivelmente, isso se deve à falta de

experiência no uso de tecnologias móveis e não por problemas com o protótipo.

Também houve relato de que algumas imagens estavam com baixa resolução, sendo difícil compreendê-las. Como sugestão, os estudantes relataram que poderia ter um recurso para ajustar o tamanho da fonte.

Com relação à atratividade, que envolve características que possam atrair um potencial usuário para o sistema. Identificou-se que podem ser adicionados novos recursos, principalmente com relação a compartilhamento de informações.

## 5. Considerações

Uma das vantagens de aplicativos em relação ao formato de arquivo ePub é que em um aplicativo pode-se explorar a interatividade de diversas formas, pois se trata de um software desenvolvido para aquele fim. O conteúdo dos cadernos didáticos no formato de aplicativos é ideal para livros que devem ser usados pelo usuário como um guia. Ao final, o conteúdo estará em um software, portanto, segue uma forma de produção diferente dos ebooks.

Existem diferentes e interessantes iniciativas de pesquisa no que se refere a explorar a leitura e o aprendizado móvel. Por exemplo, os materiais elaborados para ensino a distância, que em sua maioria são em formato PDF, também são acessíveis a estudantes por meio de um dispositivo móvel como *tablets* ou *smartphones*. Porém, utilizar tais dispositivos apenas para acessar um conteúdo que se apresenta em sua maioria de forma estática e com interatividade limitada ou inexistente, é deixar de explorar os potenciais inerentes à mobilidade.

Dessa forma, o protótipo desenvolvido é uma tentativa de explorar a leitura e a educação no contento da mobilidade. Ainda que em fase inicial e com uma avaliação superficial, salienta-se conforme análise dos resultados preliminares que adaptar material didático para dispositivos

móveis é uma tarefa complexa, que requer esforço de pesquisa, mas que pode ser recompensada pelos benefícios gerados ao aprendizado.

#### **REFERÊNCIAS**

ABIRESEARCH. 2.1 Billion HTML5 Browsers on Mobile Devices by 2016 says ABI Research, 2011. Disponível em: http://www.abiresearch.com/press/3730-2.1+Billion+HTML5+Browsers+on+Mobile+Devices+by+2016+says+ABI+Resear ch. Acesso em 30 jul 2013.

ABNT - ASSOCIAÇÃO BRASILEIRA DE NORMAS TÉCNICAS. *NBRISO/IEC9126-1*. Engenharia de software - Qualidade de produto - Parte 1: Modelo de qualidade. 2003.

ABOLFAZLI, S., Sanaei, Z., Ahmed, E., Gani, A., &Buyya, R. Cloud-Based *Augmentation for Mobile Devices*: Motivation, Taxonomies, and Open Challenges. 2013. arXiv preprint arXiv:1306.4956.

ALLEN, Sarah; Graupera, Vidal; Lundrigan, Lee. Phone Gap. In: Pro Smartphone Cross-Platform Development. *Apress*, 2010. p. 131-152.

ALLY, Mohamed. (Ed.) *Mobile Learning*: Transforming the Delivery of Education and Training. 2009.

CSS3. CSS Cascading and Inheritance Level 3. W3C Last Call Working Draft 30 July 2013. Disponível em: http://www.w3.org/TR/2013/WD-css-cascade-3-20130730/ Acesso em: ago 2013.

DUARTE, A. F.; Duarte Filho, N. F.; Barbosa, E. F.; Costa, N. P. O. Protótipo de Apoio ao Ensino e Aprendizagem de Endometriose por meio de Dispositivos Móveis. In: **Anais** do XXIV Simpósio Brasileiro de Informática na Educação. 2013.

GOULARTE, Fábio Bif; Wilges, Beatriz; Nassar, Silvia Modesto. Uma Proposta de Material Didático Segundo as Características do m-learning. *RENOTE*, v. 11, n. 3, 2013.

JUNTUNEN, Antero; Jalonen, Eetu; Luukkainen, Sakari. HTML 5 in Mobile Devices - Drivers and Restraints. In: System Sciences (HICSS), 2013 46th Hawaii International Conference on. IEEE, 2013. p. 1053-1062.

MANDULA, Kumar; Medat, Srinivasa Rao; Muralidharan.V; Parupalli, Ramu. A Student Centric Approach for Mobile Learning Video Content Development and Instruction Design. 2013 January 27 - 30. 2013 ICACT2013.

MORAIS, Felipe de; Jaques, Patrícia Augustin. PATtoPhysics - Um resolvedor de equações algébricas como ferramenta de apoio à sala de aula no ensino de

Movimento Retilíneo Uniforme. 2013. In: *LACLO* 2013, Valvidia, Chile, Outubro de 2013.

MÜHLBEIER, A. R. K., Mozzaquatro, P. M., Medina, R. D., de Oliveira, L. C., Moreira, R. C., & Antoniazzi, R. L. MOBILE HQ: O Uso de Softwares Educativos na Modalidade M-Learning. In: *Anais* do Simpósio Brasileiro de Informática na Educação. 2012.

PALMIERI, Manuel; Singh, Inderjeet; Cicchetti, Antonio.Comparison of cross-platform mobile development tools. In: Intelligence in Next Generation Networks (ICIN), 2012 16th International Conference on. IEEE, 2012. p. 179-186. RAMADATH, S., Collins, M. Mobile Application Development: Challenges and Best Practices, 2012. Disponível em: http://www.accenture.com/SiteCollectionDocuments/PDF/Accenture-Mobile-Application-Development-Challenges-Best-Practices.pdf Acesso em: jul 2013. SACCOL, Amarolinda; Barbosa, Eliane Schlemmer Jorge. M-learning e u-learning: novas perspectivas da aprendizagem móvel e ubígua. Ed. Pearson.

SCANLON, Eileen; Jones, Ann and Waycott, Jenny (2005). Mobile technologies: prospects for their use in learning in informal science settings. *Journal of Interactive Media in Education*, 21(5), 1-17

SHAH, Neil. *One Billion HTML5 Phones to be Sold Worldwide in 2013*, Strategy Analytics, 2011. Disponível em: http://www.strategyanalytics.com/default.aspx?mod=pressreleaseviewer&ao=5145 Acesso em: 30 jul 2013.

SINCLAIR, G.; Mcclaren, M.; Griffin, M.J. *E-learning and beyond*. Victoria, BC: British Columbia Ministry of Advanced Education, 2006.

TIM O'Reilly. What Is Web 2.o. 2005. Disponível em: <a href="http://oreilly.com/web2/archive/what-is-web-20.html">http://oreilly.com/web2/archive/what-is-web-20.html</a>. Acesso em: 15 jul. 2013.

W<sub>3</sub>C. Web Design and Applications: JavaScript Web APIs, 2013. Disponível em: http://www.w<sub>3</sub>.org/standards/webdesign/script Acesso em: jul 2013.

WILGES, Beatriz et al. Especificação de uma ontologia de convergência entre a formação e atuação dos professores do Sistema e-TEC. *RENOTE*, v. 9, n. 2, 2012.

WOODS, Will; Scanlon, Eileen. iSpot Mobile - A Natural History Participatory Science Application. 2012. In: *Proceedings of Mlearn 2012*, Helsinki, Finland, Oct 15-16.

# PARTE III ESTATÍSTICA TEXTUAL

# Metodologia, ciências duras e literatura em meio digital Saulo Cunha de Serpa Brandão<sup>1</sup>

Inicio a escrita deste texto com uma preocupação que divido já com meus leitores, de forma que eles possam decidir se continuam a leitura, se o tempo vale a pena. O alerta vem por conta de eu estar assumindo riscos próprios de quem se aventura a escrever ensaios. Charles Lamb (Elia) dizia que os escritores de ensaios têm um tipo de *imperfect intellects*, ele mesmo sendo assim, segundo o próprio, que os libera da necessidade de serem muito precisos. Seu quase-contemporâneo Alexander Smith apregoava que os ensaístas eram "revisores libertinos", obedecendo regras próprias, só deles. Mas essas pretensas imperfeições de defensores do ensaio escondem uma realidade insofismável: conhecimento sério é construído às expensas de ensaios². A maior imperfeição deste meu trabalho está no fato de eu estar o escrevendo como um ensaio, que é por natureza instável, para reclamar um método sistemático para quem lida com literatura em meio digital. É um paradoxo.

Nós, humanistas, que nos aventuramos a conduzir estudos na área de Humanidades Digitais temos que nos preocupar com a forma que nossas pesquisas são retratadas em relatórios, artigos e outras produções técnicas. Isto porque, desde a antiguidade grega, filósofos

<sup>&</sup>lt;sup>1</sup> Professor do Programa de Pós-Graduação em Letras da Universidade Federal do Piauí (PPGL/UFPI), coordenador do Núcleo de Pesquisa em Literatura Digitalizada (NUPLID/UFPI) e do *English Research and Information Center* (ERIC/UFPI). Contato: saulo@ufpi.edu.br

<sup>&</sup>lt;sup>2</sup> Informações sobre ensaio retirados de Klaus, C. e Ned Stukey-French, *Essayist on the essay*, Iowa City, Univ. of Iowa P., 2012.

vêm aprimorando (modificando) métodos para que a ciência, especialmente a pesquisa científica, seja entendida, conduzida e explicada. No mais das vezes, esses pensadores questionam as técnicas empregadas para examinar o fenômeno e obter conhecimento desta prática.

Essa discussão e construção de metodologia vêm acontecendo desde Aristóteles (talvez antes), passando por Ptolomeu, Roger Bacon, William Ockham, Francis Bacon, Isac Newton, Whewell, Einstein, Peirce, Popper, Bergson, Thomas Kuhn, toda uma geração de lógicopositivistas (Ciclo de Viena) e muitos outros. Com isso, eu não quero dizer que esses pensadores estavam intelectualmente alinhados e construindo um objeto único. Não, eles estavam, quase sempre, em desacordo, especialmente quanto às propriedades do objeto ou fenômeno observado, como este se relaciona com nossos aparatos cognitivos e como a mente age na relação com o objeto: intuitivamente, dedutivamente, absolutamente, relativamente, empiricamente etc.

A discussão sobre a natureza do objeto, ou do fenômeno, e sua relação com a nossa observação e como exploramos o objeto, como mencionado acima, vêm construindo um modelo que, hoje, serve para a pesquisa científica e no ventre da comunidade científica existe um propósito explícito e forte de que o modelo deve ser seguido, preservado e defendido. O modelo estabelecido pode (e deve, de acordo com a sociedade do conhecimento) ser adotado por qualquer que gueira que pesquisa seja sua Independentemente da posição filosófica do cientista, seja ele realista (em qualquer das 17 possibilidades de realismo que existem hoje), positivista, lógico-positivista, relativista, estruturalista.

Depois de séculos discutindo e filosofando sobre o assunto, concordando com uns discordando de outros tantos, os filósofos da ciência (na maior parte das vezes os próprios cientistas se incumbiam dessa tarefa) chegaram a um modelo que hoje é plenamente aceito e conhecido como Método Científico para pesquisa, ei-lo:

- 1- Formulação de uma pergunta (orientadora da pesquisa): nesse momento o pesquisador deve estar lendo outras experiências sobre o mesmo objeto ou sobre uma metodologia aplicada em uma pesquisa similar a sua. Analisando as falhas e acertos de outras pesquisas, consultando outros pesquisadores etc. É uma etapa fundamental para a boa orientação e condução da pesquisa, além de servir de norte para que o cientista não se afaste de seu objetivo.
- 2- Definição da(s) hipótese(s) (também conhecida por conjectura[s]): com o conhecimento adquirido durante a formulação da pergunta orientadora, o cientista pode propor hipóteses restritas ou abertas. O pesquisador sempre peleja contra a hipótese nula ou falsa.
- 3- Predição(ões): com a(s) hipótese(s) formulada(s), o cientista parte para o exercício de predizer o que vai acontecer na pesquisa. As duas predições mais apreciadas são: a predição se realizou ou o resultado do experimento ainda não é conhecido.
- 4- Teste: As hipóteses são testadas experimentalmente. Há pesquisadores que são tão altivos que se a hipótese não se realizar, eles alegam que os técnicos do laboratório erraram na condução do teste.
- 5- Análise dos resultados.
- 6- Tirar as conclusões.
- 7- Publicar os resultados.

Existe um último passo para validar a pesquisa e seus resultados, mas que não pode ser realizado pelo mesmo cientista que conduziu o experimento. Trata-se da replicação do experimento por outro cientista em outro laboratório, obtendo o mesmo resultado. Essa etapa é um dos principais axiomas da ciência moderna, mais do que um axioma, uma pilastra de sustentação de toda a ciência e metodologia científica. Dada

a importância dessa etapa, o cientista tem que incorporar o bom exercício de registrar cada passo da pesquisa com detalhamento de um ourives, nos mínimos filigranas.

Feito esse resumo da situação da pesquisa científica em relação a uma metodologia da ciência, passemos agora a um rápido cenário da pesquisa nas Humanidades e em especial na Literatura e, não poderia deixar de falar um pouco, da História. Escolho essas duas pelo fato de que uma, a última, esteve em uma situação muito confortável quanto ao seu método de pesquisa e já não está mais. A primeira, porque nunca encontrou uma fórmula ou modo que se prestasse a um estudo metódico de seu objeto. Obviamente, pela natureza do objeto estudado na Literatura, que é construído utilizando signos que são polissêmicos e que cada autor tem um estilo próprio de escrita e mais questões sobre a liberdade de interpretação, não há como se exigir a prova de replicabilidade que valida as Ciências Duras, com as pesquisas do tipo tradicional<sup>4</sup>.

Falemos, logo, e rapidamente, sobre a última. A História entrou em uma crise sem precedente e sem solução previsível. Ela que era tida como uma das disciplinas das humanidades que tinha um protocolo de pesquisa dos mais bem definidos e estáveis. Mas esses protocolos não resistiram a novas posições filosóficas assumidas no início do século XX. Acredito que os estudos da História começaram a ser puídos com a publicação de Ortega y Gasset, de 1928, "Filosofa de la História' de Hegel a Historiologia". Neste livro o pensador espanhol tratou de ditar o que seria uma epistemologia dos estudos da História, a Historiologia. Esta tocou em pontos que antes eram inquestionáveis para historiadores, ou havia uma quase-unanimidade entre eles sobre a sistematização dos estudos da História, uma vez que começou a discutir questões envolvidas com a narração e interpretação.

\_

<sup>&</sup>lt;sup>4</sup> Sobre pesquisas nas Humanidades e na Literatura, vale, de novo, minha vivência e leituras acumuladas e Markievicz, H., *Theory and method in literary studies*, Organon, 28 30, 1999-2001, p. 161.

Em movimentos posteriores, historiadores começaram a questionar o ambiente cultural no qual o narrador está inserido para, daí, tentar entender melhor como ele está fazendo a interpretação do real. Esse questionamento faz todo sentido e foi importante na época que foi defendido por Roger Chartier, na década de 80 do século XX. Mas com o avanço dos Estudos Culturais e a rapidez (pressa) que alguns pesquisadores dessa linha se apropriaram das ideias de Chartier, o narrador foi colocado em xeque, diria até em um xeque-mate. Pois ele passou muito rapidamente a ser considerado um informante que sistematicamente falseia o real. Daí fica a questão enfrentada pelos pensadores da História, hoje: confiar em quem? De repente toda a História foi tomada por uma cisma, tudo é movediço e estranhamente relativo. O narrador não é mais confiável.

Para complicar ainda mais o ambiente, alguns poucos pesquisadores do passado, e que hoje são uns muitos, passaram a tomar a Literatura (a arte) como sendo fonte segura (muitas vezes como fonte primária [vários artigos da área que tentam recriar o ambiente cultural ou social do Rio de Janeiro com informações pesquisadas em contos e romances de Machado de Assis, por exemplo]) para o estudo da situação do homem no tempo. Não adiantou nada Fernando Pessoa e seus admiradores cantarem aos 4 ventos que "O poeta é um fingidor ...". Para a História, ou para alguns historiadores, o poeta é o grande cronista do real (eu até defendo esta posição, mas a partir de um estudo em que a metáfora é uma ferramenta de cosmogônica [isso praticado por poetas fortes, como quer Harold Bloom]). Resumindo, as coisas, em matéria de sistematização da pesquisa na área de História, tendem a azedar ainda mais com o passar do tempo, principalmente porque, a academia está formando profissionais dentro desse modelo de pensamento que, a meu ver, não traz qualquer estabilidade para as pesquisas na área.

Na Literatura, o estado da arte em relação à metodologia da pesquisa é sofrível. Na realidade, os teóricos e filósofos das artes (Literatura) nunca chegaram a um acordo sobre uma metodologia adequada para o estudo

do texto literário, seja ele em prosa ou verso. Mesmo sem querer aborrecer meu leitor com citações, não tenho como passar sem referenciar essa passagem de Henryk Markiewicz, observem o que o pensador polonês (já citado) diz nas primeiras linhas de um artigo sobre História e Teoria da Literatura:

A scholar who appreciates order and accuracy will be upset by the profusion, as well as the ambiguity and interference, of notions and terms relating to the theoretical branch of the science of literature. Even poetics and rhetoric, both descended from the Aristotelian tradition, combined normative with descriptive aims, overlapped each other by their scopes (think of the theory of tropes and figures)...

O pensador mais adiante em seu artigo vai dizer que existe uma tentativa grande por parte de alguns filólogos (neste ensaio, entenda-se o estudioso das línguas e literaturas) em sistematizar os estudos da Literatura e cita Wilhelm Wackemagel, que foi um dos pioneiros no estudo das ciências das artes. Ele classificou o estudo da Literatura da seguinte forma: o estudo da natureza e tipos de poesia deve ser o trabalho da poética; a literatura em prosa deveria ficar sob a reponsabilidade dos retóricos; e, o estudo da linguagem tanto na poesia como na prosa seria desenvolvido pelos estilistas. Mas isso foi no século XIX

Se os humanistas não conseguiam, nem conseguem, desenvolver uma metodologia dentro da área, intrinsicamente, tão pouco alargavam um relacionamento amistoso com as outras áreas. São famosas as disputas entre humanistas e cientistas no curso do tempo<sup>5</sup>.

Os estudiosos da área sequer se entendem sobre qual nome devem usar para a ciência, se: Teoria da Literatura, Teoria Literária, Ciência da

<sup>&</sup>lt;sup>5</sup> Mais sobre esse assunto pode ser visto em um artigo de minha autoria publicado na Texto Digital, de título *As duas culturas revisitadas – de novo* (https://periodicos.ufsc.br/index.php/textodigital/article/view/18079288. 2013v9n2p230 /26061).

Literatura, Estudos Literários, Filosofia da Literatura, Literatura etc. Existem diferenças estabelecidas entre alguns dos termos, mas, o mais comum é que os estudiosos nem saibam disso e terminam usando um pelo outro, alhos por bugalhos.

Vale lembrar que no século XIX ainda valia a física moderna e nesta a teoria estava completa, uma vez que respondia a todas (quase todas, na realidade) as questões sem ferir a lógica do conhecimento da época. Todos os fenômenos conhecidos vinham já acompanhados de bulas que explicavam cabalmente o fenômeno e em maior ou menor escala proclamavam a supremacia da ciência. Valia ainda, naquela época, a metáfora do relógio para se referir ao universo. Tudo funcionava com a precisão de um Patek Phillipe. Mesmo assim, o estudo da Literatura ainda não conseguia se organizar.

O século XX veio para complicar ainda mais esse panorama. Vejamos no Brasil como a arte evoluiu. Com o experimentalismo ácido do primeiro Modernismo, o literato ficou ainda mais entrecortado, porque o pouco que existia de sistemático, vindo do século anterior, se estraçalhou completamente. O pouco de aparato teórico comungado pela maioria de teóricos e críticos já não conseguia dar conta das inovações que viraram norma na Literatura nas primeiras décadas do XX. Mas o foco dos estudos literários ainda apontava prioritariamente para o texto artístico em prosa ou verso.

Com a chegada da Segunda Geração a coisa desandou de vez. Com o aparecimento dos romances regionais, mais politizados e com forte carga do drama social, o literato assumiu de vez a sua face fragmentada, virou sociólogo, historiador, antropólogo, geógrafo, cientista político, jurista, filósofo, ou seja, ele era o sujeito das Humanidades, entendia de tudo... e de nada. Porque ele não é profissional de nenhuma das profissões citadas, e, pior, esqueceu de estudar Literatura. É um pobre coitado, é uma caricatura de pesquisador.

Esse era o panorama dos estudos literários até o fim da Segunda Grande Guerra. Após 1945 o caldo entornou mais ainda, o que era mal, ficou pior. Isto porque chegava uma nova leva de artistas muito preocupada com a língua, investigando e manipulando a língua portuguesa, fazendo com que a língua investigasse a própria língua e elas tecessem os limites antes respeitados: era a 3ª Geração de Modernistas.

Esse panorama artístico complexo tinha como pano de fundo a reviravolta científica, em que as certezas da Física Moderna foram desmanteladas pelos estudos da natureza do mundo subatômico. As partículas, de repente, se tornaram rebeldes e só aceitavam mudanças diante de um pacote de energia, era quase uma organização sindical. Nenhuma agia sozinha, ou à troca de pouca coisa. Ou mudavam todas, ou não mudava nenhuma e ainda mediante um pacote de incentivos (energia). Era a chegada da Mecânica Quântica: Heisenberg, Du Bois, Einstein, Plank, Bohr, Schrödinger, Born, Feynman, e mais vários outros.

Outras teorias e fenômenos novos apareceram para revolucionar o sistema de conhecimento humano, como: Teoria das Cordas, Princípio da Incerteza, Caos como elemento criador, Teorias da Relatividade, Modelo de Rutherford-Bohr, Buracos de Minhoca, Fótons Gêmeos, Gato de Schrödinger, Buracos Negros, Teoria do Big-Bang.

Toda essa parafernália científica estava acontecendo e sendo socializada nos primeiros 60 anos do século XX. E, claro, o povo das Humanidades e das Sociais não podia ficar só de espectador. Ele tinha que meter a mão nessa nova ciência. Começou-se a comparar os modelos subatômicos, incerteza, caos, gatos, gêmeos com o comportamento do homem em sociedade e consequentemente, como essa nova fisionomia aparece em romances e poesias. Com isso apareceram os: Baudrillard, Morin, Huxley, Evans, Jeffares, Snow, Hayles, Saulo Brandão (eu mesmo). Alguns com muito cuidado e respeito às diferenças (me incluo neste grupo), mas outros de forma

mais bombástica e buscando chamar atenção da mídia especializada. E este comportamento fez com que Alan Sokal, físico da Universidade de Nova York, publicasse um artigo *non-sense*, na revista *Social Text* (Duke University), onde fingia argumentações científicas para sustentar as verdades ideológicas dos editores da revista (visada de esquerda e pósmoderna). E, simultaneamente, publicou outro artigo na revista *Lingua Franca* onde revelava a fraude intelectual e o descompromisso intelectual dos editores da *Social Text*. Isto tudo em 1996.

A conclusão desse panorama social/intelectual/artístico mal-amanhado do fim do século XX e começo do XXI é que a pesquisa na Literatura tomou um caráter extremamente teorizado, mas nem sempre com teorias próprias da área, mas de textos tomados de empréstimo das Ciências Sociais e de disciplinas das Humanidades. Como parecerista de algumas revistas acadêmicas, recebo com certa frequência artigos para emitir parecer e, com base em uma observação empírica, mas sem a devida coleta de dados para estudos estatísticos, posso arriscar dizer que entre 60 e 75% da trama dos artigos serve apenas para que o articulista procure se auto-referendar como conhecedor de alguma teoria exógena para justificar o fato de usá-la para analisar o texto literário. O resultado dessa prática é, muitas vezes, artigos à la Sokal, citado mais cedo neste texto, sendo que com bom propósito. Os restantes 25 a 40% do texto é dividido entre introdução, conclusão e um pouquinho de reflexão sobre o texto literário.

Existe, também, obviamente, aqueles que não se cercam de teorias exógenas e fazem a análise de forma subjetiva e pouco convincente. Tratam o texto como faziam os seguidores da escola do *Close Reading*. Esta fez história nos Estados Unidos dos anos 50 do século XX e também na França, onde era tratado como *Explication de Texte*. Nessa visada, analisa-se uma porção pequena do texto e trabalha-se para persuadir o leitor de que aquele viés serve para o texto inteiro.

Obviamente, existem muitos pesquisadores escrevendo artigos muito bem estruturados, bem fundamentados. Utilizando teorias, poéticas, estilísticas próprias da área da Literatura, ou se cercando muito cuidadosamente de aparatos de outras artes. Fazendo o trabalho de um real pesquisador da área. Mas, certamente, não são maioria.

Existem pesquisadores que justificam a utilização de teorias exógenas ao fato de não existirem estudos suficientemente sofisticados para dar conta da complexidade da produção literária contemporânea (me refiro ao Brasil e à Literatura Brasileira). E isto pode até ser verdade, mas a atitude mais coerente seria começar a teorizar imediatamente, pois é evidente o sinal que existe esse nicho de mercado para aqueles que estejam dispostos a pensar sobre o estado da arte. No entanto, não me ocorrem nomes de estudiosos que estejam presentemente pensando Literatura enquanto objeto de pesquisa. Temos uma mancheia de pensadores na geração passada, mas que em seus dias já eram reconhecidos e aclamados. Digo isso para arriscar dizer que se os pensadores existissem hoje, eles provavelmente seriam também reconhecidos. Quando falo de gerações passadas me refiro aos irmãos Campos, Pignatari, Philadelpho Menezes, Carpeaux, Amoroso Lima, Mario Faustino e mais alguns outros.

Outro problema grave que as Ciências Duras já resolveram e nós, das Humanidades, nem tanto é a competência linguística para lidar com línguas estrangeiras, especialmente o inglês. O pesquisador das Ciências Duras que não lê textos em inglês está fadado ao fracasso. As melhores revistas brasileiras só aceitam textos em inglês. Poderia citar dezenas de exemplos, mas vou me ater a um, que tal? *Brazilian Journal of Medical and Biological Research*. Isto dito, volto a questionar como a Literatura e as Humanidades, especialmente no Brasil, podem evoluir se elas não dominam (leitura e escrita, pelo menos) uma língua estrangeira representativa para a ciência? Existem pensadores produzindo muita teoria, mas em inglês/francês/espanhol. Exemplos: N. Katherine Hayles, Fredric Jameson, Homi Bhabha, Kristeva, Spivak, Roberto Echevarría, Umberto Eco, etc.

Esse panorama que tracei do estudo de Literatura no Brasil pode ser pessimista, eu carrego um pouco dessa pecha comigo, mas o leitor esperto pode tirar uma prova dos nove do que eu escrevo e perceber um panorama muito próximo do real.

Tudo isso dito, fica revelada a minha visão do passado/presente da sistematização da pesquisa nas várias áreas de pesquisa, em especial nas Humanidades e na Literatura. Passarei, a partir deste parágrafo a tratar do presente/futuro. Tratarei da sistematização nas Humanidades Digitais/Literatura em meio digital. E não adianta dizer que isso é o futuro! Isso é o presente e o futuro, cada vez mais. Mas dentro dessa área vou me apegar ao meu interesse primeiro: Estilometria e Lexicometria, as quais já vi alguém tratando do assunto de uma forma mais geral: textometria.

O estudo dos textos através das contagens numéricas das palavras é muito antigo. Faz parte da tradição filológica do Cristianismo. Estudiosos dos textos bíblicos já faziam contagens de textos no século XVIII. No mundo acadêmico, o pioneiro em estudos estilométricos foi Agustus de Morgan, matemático inglês que em 1851 propôs que as dúvidas tocantes à autoria dos textos de São Paulo fossem superadas medindo o comprimento, número de letras utilizadas nas palavras nas diversas epístolas em disputa. Mas ele não passou da formulação da proposta. Foi T. C. Mendenhall quem primeiro aplicou uma análise espectroscópica. Utilizando a prática, ele mostrou que o gráfico resultante da aplicação da regra nas obras de Charles Dickens apresentava um resultado uniforme e diferente de diversos outros autores. Isto em 1887. Desde então, com altos e baixos, as análises numéricas de textos literários vêm aparecendo em revistas especializadas.

Com o aparecimento dos computadores no século XX, especialmente a partir da popularização dessas máquinas com a criação dos *Personal Computers*, na década de 1960, os humanistas passaram a conviver mais

proximamente aos computadores e a consequência natural foi o maior uso deles para explorar seus objetos de pesquisas.

Na área das Letras, os linguistas, como sempre, se estruturaram mais rapidamente e avançaram com os estudos de corpora, que já eram praticados na Grécia antiga, conduzidos de forma manual. Em 1964, apareceu o *Corpus Brown*, que foi o primeiro corpus eletrônico e contava com 1 milhão de palavras em língua inglesa. Hoje, a Linguística de Corpus é um ramo da Linguística completamente estruturado e documentado. Conta com metodologia própria que muito se assemelha à metodologia utilizada pelas Ciências Duras. E existe uma grande incidência desses estudos no ensino de línguas estrangeiras e também pelos dicionaristas.

Na Literatura, os profissionais vêm se dedicando a alguns vieses que se apresentam como mais promissores e de resultados mais impactantes. Eles têm trabalhado com leitura de hipertexto, recuperação de documentos literários, recuperação de obras, disponibilização de obras na internet e estudos estatísticos de obras literárias. Destes listados, existem variações que acontecem de acordo com o propósito de cada pesquisador, ou grupo de pesquisadores.

Em qualquer das variáveis, os pesquisadores devem ter em mente a necessidade de registrar cada passo de sua pesquisa com a preocupação que eles estão entrando em uma nova seara. Os artigos científicos, para esses indivíduos, deixam de ser aqueles com tendências fortemente impressionistas e largamente fundadas em teorias, para adentrar em um novo modelo, no qual, via de regra, as citações remetem para resultados de outras pesquisas sobre objetos semelhantes, senão os mesmos, que servirão como parte da nova pesquisa e/ou para direcionar a nova pesquisa para um objetivo diferente das anteriores. Nesse novo panorama, os artigos de Literatura devem responder ao quesito de repetitividade que norteia as Ciências Duras. Ou seja: os resultados obtidos por um pesquisador devem ser repetidos por outro e, seguindo a metodologia registrada pelo primeiro, chegar aos mesmos resultados.

Destaco aqui um exemplo dessa prática científica já no século XIX. Lewis Campbell resolveu a ordem cronológica dos diálogos de Platão em 1867. Trinta anos depois, Ritter, filólogo alemão, chegou aos mesmos resultados que Campbell, seguindo os registros de pesquisa de 1867<sup>6</sup>. Essa prática de registrar o passo-a-passo da pesquisa em uma parte do artigo tratando da metodologia passa a ser um *must* para os pesquisadores que se adentrarem nessas práticas.

O leitor mais interessado nessa visada metodológica, na problemática de desencontro em resultados de pesquisa, e querendo se aprofundar no assunto, pode encontrar na internet algumas disputas sobre experimentos realizados e não replicados seguidos as mesmas trilhas.

Os literatos vêm utilizando ferramentas telemáticas e programas computacionais cada vez mais. É relativamente fácil encontrar estudos para verificação de autoria de textos apócrifos, sob pseudônimo ou de autoria desconhecida. Por uma questão de importância, Shakespeare é o autor mais examinado. Mas também encontram-se muitos estudos sobre o estilo de Charles Dickens. Este por ter um estilo muito marcante, estável, e todas as suas obras estarem disponíveis na internet. Pelo fato de sua estabilidade estilística ele é muito utilizado como *corpus* testemunha para validação de metodologias.

Os softwares utilizados são do tipo data mining, o Gartner Group define esse processo como o de descobrir novas relações significativas, padrões e tendências distribuídas em grandes corpora usando tecnologias de identificação de padrões e técnicas estatísticas e matemáticas. Essas técnicas têm como características: demonstrar as relações internas do texto de forma organizada; e, servir a estudos interdisciplinares.

Os projetos de análise de grandes *corpora* utilizando *data mining softwares* (DMS) devem ser estruturados em 6 partes, a saber: 1 –

-

<sup>&</sup>lt;sup>6</sup> Kenny, A., The computation of style, NY: Pergamon P., 1981, p. 3.

Reconhecimento do problema e projeto; 2 – Análise e entendimento dos dados; 3 – Preparação dos dados; 4 – Fase de modelagem dos dados (organização dos dados para otimização dos resultados); 5 – Fase de avaliação (verificar se a modelagem dos dados atingiram os resultados); e, 6 – Fase de melhoria dos padrões (checar se obteve-se o melhor resultado para daí voltar para a parte 3, se necessário). O loop descrito para a parte 6 pode também ser necessário na parte 5<sup>7</sup>.

Da mesma forma que existem 6 fases na preparação para a análise de dados, existem também 6 falácias facilmente encontradas em trabalhos com DMS, elas são: 1 – Existem softwares que trabalham livremente os dados e no final apresentam os resultados; 2 – Os softwares não precisam de intervenção humana para analisar os dados e dar a solução para os problemas; 3 – DMS retornam o investimento feito na compra deles rapidamente; 4 – DMS são intuitivos e fáceis de usar; 5 – DMS identificarão o problema de seu dados e sua pesquisa; e, 6 – DMS limparão seus dados automaticamente. Sobre este último, limpeza de dados, o literato (ou qualquer profissional) tem que ter muito cuidado, porque um corpus mal trabalhado, mal limpo, mal organizado, fatalmente arruinará o projeto e a pesquisa.

Continuando com o número 6, eles são os resultados que deve-se esperar de uma análise com DMS: 1 – Exposição de padrões dentro do corpus; 2 – Permitir uma estimativa (dados numéricos); 3 – Predição (permite estimar variações dos padrões no futuro); 4 – Classificação dos dados; 5 – Agrupamento de certos elementos dentro de uma lógica; e, 6 – Definição de associações por atribuições próprias de certos elementos.

Retorno à necessidade de pré-processar todo o *corpus*. O investigador tem que lidar com a fase de preparação do material cru. Essa fase é muito trabalhosa e não pode estar errada. Um erro ou desleixo na

<sup>&</sup>lt;sup>7</sup> Mais informações podem ser obtidos em http://www.nancygrady.info/CRISP-DM.pdf (as informações reunidas aqui vieram desse *site*).

preparação do corpus acarretará em resultados errôneos para toda a pesquisa. Dependendo da complexidade e tamanho do corpus, essa fase da pesquisa pode consumir 60% do tempo de execução do projeto. Se o corpus estiver disponível em formato TXT na internet, projeto Gutenberg (www.gutenberg.org), por exemplo, isso poupará muito tempo ao pesquisador. Mas se este tiver que digitalizar (scanner) o texto objeto da pesquisa, para posteriormente submetê-lo a um software que transforme em DOC (Microsoft Word) ou TXT (Plain Text), geralmente softwares com base na tecnologia OCR (Optical Character Recognition), o tempo vai ser fator determinante para a condução bem sucedida da pesquisa. O processo para chegar do texto (objeto sólido, livro, jornal) ao formato TXT, que é exigido pela maioria dos DMS, é extremamente trabalhoso. Gera um turbilhão de erros e o pesquisador tem que fazer a limpeza cotejando palavra por palavra (texto original e texto transformado para TXT), especialmente de textos literários de personalidade vanguardistas, em que o autor cria neologismos, usa palavras em outros idiomas, arcadismos etc. Existem softwares para fazer a limpeza dos corpora, mas eles tão pouco são confiáveis.

Com essa explanação pretendemos chamar a atenção de novos pesquisadores que intencionam adentrar no mundo das Humanidades Digitais, especificamente da Literatura, que existe já uma massa crítica razoável e acumulada há muitos séculos, que essa tradição exige uma metodologia mais comprometida com a replicação da pesquisa e como pode ser trabalhoso fazer pesquisa na intersecção da Literatura com a Ciência da Computação. É um campo muito promissor e faz com que os pesquisadores humanistas sejam valorizados no mundo da cultura, em especial no contexto de preservação de obras literárias, cartas, jornais, originais etc. Outras áreas, como os geneticistas literários, ganham com a disponibilidade de documentos originais e possíveis de serem acessados de qualquer lugar que tenha um ponto de internet. No estágio em que estamos já é muito frustrante quando descobrimos que

<sup>&</sup>lt;sup>8</sup> Pyle, D., *Data Preparation for Data Mining*\_San Francisco, Morgan Kaufmann Publisher, 1999.

algum corpus não está ainda *on-line*. São estudos promissores, novas verdades serão reveladas por conta dessa nova visada nos estudos literários, mas existe muito trabalho pela frente até que as promessas virem realidades.

# Métricas rosianas (1929-1967)

Verônica Ribas Cúrcio<sup>1</sup>

Neste capítulo apresentaremos alguns resultados de um estudo de estatística textual realizado sobre a obra de João Guimarães Rosa. Vale dizer que dois campos foram explorados: o primeiro campo retrata as características gerais e específicas do vocabulário de Guimarães Rosa e o segundo compara os resultados com três teses da crítica literária rosiana.

Os textos que compuseram o corpus foram: *O mistério de Highmore Hall* (1929); *Makiné* (1930); *Tempo e destino* (1930); *Caçadores de camurças* (1930); *Magma* (1936); *Sagarana* (1946); *Corpo de baile* (1956); *Grande sertão: veredas* (1956); *O último dos maçaricos* (1958)²; *O mistério dos MMM* (1961); *Primeiras estórias* (1962); *Tutaméia* (1967); *Estas estórias* (1969) e *Ave, palavra* (1970)³.

Reunimos de modo abrangente o léxico (1.094.481 ocorrências, 58.647 vocábulos) de Guimarães Rosa dentro de um percurso cronológico de sua produção literária, utilizando a ferramenta informatizada chamada *Hyperbase* (desenvolvida pelo laboratório BCL da Universidade de Nice),

<sup>&</sup>lt;sup>1</sup> Pesquisadora do Núcleo de Pesquisa em Informática, Literatura e Linguística da Universidade Federal de Santa Catarina (NuPILL/UFSC). CONTATO: demodoco@gmail.com

<sup>&</sup>lt;sup>2</sup> Tradução condensada do inglês, da obra *The Last of the Curlews* (1954), de Fred Bodsworth.

<sup>&</sup>lt;sup>3</sup> As datas aqui apresentadas são as de primeira publicação, contudo, algumas dessas obras são coletâneas de textos que foram publicados em datas anteriores.

e por meio dela, fizemos nossas análises sobre os dados estatísticos dos textos.

Dos resultados, podemos destacar que mais da metade do vocabulário (52,81%) não se repete<sup>4</sup>, o que demonstra bastante riqueza e diversificação. Os textos mais curtos representaram maior riqueza lexical, os textos com mais diálogos apresentaram menor diversificação. Concluímos que a maioria das obras apresenta mais déficit de vocabulário que excedente, principalmente as que retratam a temática do sertão (Sagarana, Grande sertão: veredas e Corpo de baile).

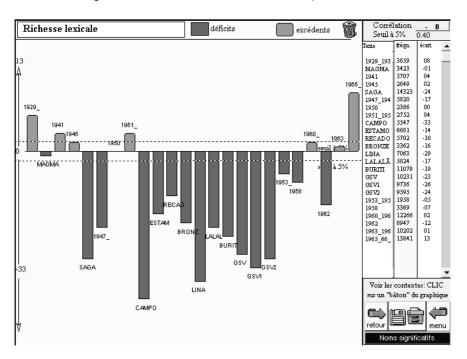


Gráfico 1 – Riqueza lexical da obra de Guimarães Rosa

<sup>&</sup>lt;sup>4</sup> Por conta das análises de tempos verbais, optamos por não lematizar o *corpus*.

Os resultados dos hapaxes também nos levaram à mesma conclusão: Ave, palavra surge como a obra mais diversificada. Das análises dos resultados, vimos que a riqueza lexical é avaliada a partir de critérios como gênero literário, estilo e temática, contudo, no caso de Rosa, o tema não parece contribuir com a diversificação do léxico. Sabemos que o tema sertanejo é bastante presente nas obras *Tutaméia*, *Estas estórias e Primeiras estórias*, porém, a ênfase no vocabulário é dada nas três obras destacadas (*Sagarana*, *Grande sertão: veredas e Corpo de baile*). Em resumo, podemos afirmar por meio das análises de crescimento de vocabulário que, em Rosa, a riqueza vocabular não se apresenta pelo volume da obra, mas pela temática.

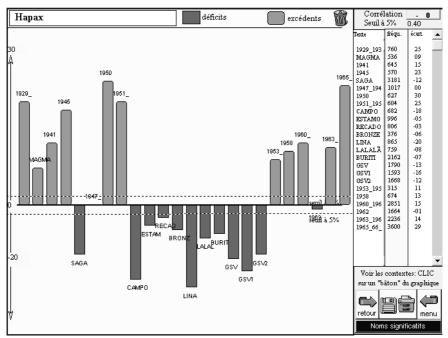


Gráfico 2 — Hapax da obra de Guimarães Rosa

### Pontuação e verbo

A pontuação rosiana também foi considerada nesse estudo. Por meio da ferramenta que trata da evolução do vocabulário, percebemos que, ao longo de sua produção, Rosa empregou mais a vírgula e o pontofinal, enquanto houve abandono da exclamação. A interrogação apresentou tendências deficitárias no início e no fim da linha cronológica, mas teve seu auge positivo nas obras *Grande sertão: veredas e Corpo de baile.* 

Das análises sobre os verbos, destacamos que os tempos verbais mais utilizados são: o subjuntivo do presente e o futuro do pretérito, por outro lado, o pretérito imperfeito foi menos empregado no final da produção.

#### Zona de isolamento lexical

Sobre os gráficos das análises fatoriais constatamos que as obras mais distantes da linha que contorna a característica rosiana foram as de gêneros distintos: os contos de um escritor principiante (1929-1930), o capítulo de um romance policial (*MMM* - 1958) desenvolvido em parceria com outros escritores e a tradução condensada de um romance canadense (1960-1961).

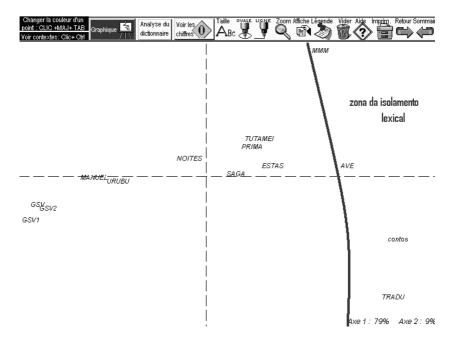


Gráfico 3 – Análise fatorial da obra de Guimarães Rosa

# Evolução do vocabulário

Dos resultados sobre os coeficientes de evolução do vocabulário, detectamos, nas altas frequências, como nos cálculos de distância lexical, mais vocábulos que se direcionam para a temática do ser e suas preocupações existenciais, tais como o *tempo*, *o amor* e a *alma*. Do mesmo modo, redefinimos a importância das palavras gramaticais, pois além das mesmas cumprirem com o funcionamento da lógica do discurso, elas também possuem carga semântica. Encontramos alguns advérbios e pronomes que identificaram dois campos semânticos: "ser" e "tempo".

Além da garimpagem dos dados e das suas descrições sobre as características gerais e específicas do vocabulário rosiano, buscamos

confirmar e complementar intuições oriundas de leituras convencionais da crítica de Rosa. Para isso, nos apoiamos em Sperber (1982), Covizzi (1978) e Daniel (1968).

Suzi Sperber (1982), em Guimarães Rosa: signo e sentimento, estuda a produção de Rosa no que diz respeito à organização da linguagem sob a luz da *Textkritik*. Nesse método, ela verificou o comportamento do vocabulário e o amadurecimento da escrita do autor, estudando desde o rascunho à versão definitiva. Das confirmações da autora, buscamos trabalhar com aquelas que viabilizariam análises sobre o crescimento e evolução de vocabulário. Vimos com Sperber (1976), que no processo de estruturação da obra rosiana, o estilo de Rosa se sobressai ao tema. Isso foi encontrado, quando analisamos o resultado sobre a variação de vocabulário, o gráfico mostrou déficit nas obras que retratam o sertão. Nas outras obras (as não-rurais) houve uma forte diversificação do léxico. Concluímos que, a temática em Rosa é um elemento de estilo, contudo ela não contribui para a diversificação do léxico.

Outro estudo que escolhemos para guiar nossas análises foi a perspectiva de Lenira Covizzi<sup>5</sup> (1978), cuja tese é delinear um percurso da narrativa rosiana que vai da expressão (arte em si) à explicação (comentário, autocrítica). Para a autora, *Grande sertão: veredas e Corpo de baile* abordam universos bem elaborados do ponto de vista da linguagem e do tema e que resultaram da densidade que foi a produção de Sagarana. A partir dos textos de 1956, ocorre uma diluição do tema e em contrapartida a produção literária ganha um aumento considerável em liberdade da linguagem. Segundo Covizzi (1978), é em Primeiras estórias que se inicia o processo explicativo da obra.

Pelo levantamento de vocabulário por ocorrências de hapax, concluímos que a fase expressiva de Rosa, a qual teria o seu cume em

<sup>&</sup>lt;sup>5</sup> Vale lembrar que Covizzi (1978) considerou para a sua análise duas categorias "expressão" e "explicação". Em nossa tese, retomamos essas categorias para prosseguir com o trabalho da pesquisadora, porém, sob outro viés e alcançando a obra completa de Guimarães Rosa (pois a autora não abarcou toda a produção).

Grande sertão: veredas, não ocorreu em termos lexicais, e que a fase explicativa se utilizou de uma diversidade lexical muito maior que a fase expressiva. Por meio dos indícios verbais extraídos dos coeficientes positivos e negativos sobre a evolução do vocabulário, chegamos em acordo com Covizzi (1978), que, ao final da carreira literária, Rosa compôs mais locuções discursivas e comentários.

A respeito do percurso narrativo de Rosa, a autora conclui que a partir dos textos de 1956 há uma redução do teor regional como em *Primeiras Estórias*. Para analisar essas afirmações, fizemos um levantamento de vocabulário pelas ocorrências de *hapax*, e desse modo verificamos o comportamento de entrada de novos *hapaxes*, caracterizando o processo de criação e renovação de vocabulário.

Outro estudo que consideramos importante para as análises estilométricas foi o de Mary Lou Daniel (1968), pois suas afirmações indicam uma ruptura no léxico rosiano em duas fases, rural (com *Sagarana*, *Corpo de baile* e *Grande sertão: veredas*) e não-rural (Primeiras estórias). A autora estabelece uma progressão qualitativa (de obra a obra) que demonstra o movimento do potencial comunicativo, e indica *Primeiras estórias* como a obra madura do autor (ressaltamos também que este estudo de Daniel (1968) é contemporâneo à obra de Rosa, portanto não abrangeu a obra completa).

Como exercício quantitativo das análises de Daniel (1968), a premissa foi verificar como essa distinção entre rural e não-rural (urbano) se reflete no léxico. Para isso, utilizamos a ferramenta que mede a evolução e outra que aborda a distância lexical. Essa última, considera o vocabulário completo que compõe cada texto do corpus, sem se preocupar com a frequência dos vocábulos, mas apenas com a presença ou ausência de uma palavra. A distância lexical também verifica a conexão entre textos por meio de seus vocabulários, que podem ser comuns (permitindo a aproximação) ou não (a distância). Dos resultados obtidos, verificamos que há dois momentos de ruptura no material linguístico de Rosa, apresentando uma grande diferença de

aparecimento de *hapaxes* e riqueza lexical, ou seja, de diversidade vocabular. As obras de maior fôlego (*Sagarana*, *Grande sertão: veredas* e *Corpo de baile*) apresentaram um vocabulário mais restrito, enquanto que as outras obras mostraram um vocabulário mais variado.

Muitos dados surgiram com a utilização do *Hyperbase* aplicado aos textos de Guimarães Rosa, o que trouxemos aqui é a ínfima parte do que foi estudado. Observamos que em muitos momentos a crítica rosiana "acertou" intuitivamente, foi o que o quantitativo pôde evidenciar. A estatística textual apenas soma e complementa a leitura antes feita. Outra condição que não deve ser ignorada é a possibilidade de ampliar as hipóteses de leitura de uma ou meia dúzia de obras para a obra completa, ou seja, verificar se alguma tendência que foi de maneira intuitiva detectada por meio da leitura convencional se aplica ou não a obra completa de um escritor.

Em suma, a estilometria possibilita trabalhar com uma grande massa de textos, que aliada ainda à capacidade de memória dos computadores atuais e à potencialidade das ferramentas tecnológicas para análise de textos, traduz muitos resultados. As ferramentas disponíveis criam possibilidades e potencializam a interpretação sobre a obra literária, cabe ao estudioso elaborar suas correlações e aproveitar as variadas formas de ser ler um texto mediado pela estatística.

#### REFERÊNCIAS

COVIZZI, Lenira Marques. O insólito em Guimarães Rosa e Borges. São Paulo
Ática, 1978.
DANIEL, Mary Lou. João Guimarães Rosa: Travessia literária. Rio de Janeiro:
José Olympio, 1968.
ROSA, João Guimarães. <i>Tutaméia</i> . Rio de Janeiro: José Olympio, 1967.
Ave, palavra. Rio de Janeiro: José Olympio, 1970.
Carta a Harriet de Onís: 4 nov. 1964. [S. l.]: [s. n.], 19-?.
<i>Grande Sertão</i> : Veredas. Rio de Janeiro: José Olympio, 1974a.
<i>Primeiras estórias</i> . Rio de Janeiro: José Olympio, 1974b.

Corpo de baile. Rio de Janeiro: José Olympio, 1976a.
Estas estórias. Rio de Janeiro: José Olympio, 1976b.
Sagarana. Rio de Janeiro: José Olympio, 1980.
<i>Primeiras estórias</i> . 12. ed. Rio de Janeiro: José Olympio, 1981.
No Urubuquaquá no Pinhém. Rio de Janeiro: Nova Fronteira, 1984.
Sagarana. Rio de Janeiro: Nova Fronteira, 1993.
Ficção completa. v. 1. Rio de Janeiro: Nova Aguilar, 1995a.
Ficção completa. v. 2. Rio de Janeiro: Nova Aguilar, 1995b.
Magma. Rio de Janeiro: Nova Fronteira, 1997.
SPERBER, Suzi Frankl. Guimarães Rosa: signo e sentimento. São Paulo: Ática,
1982.

# GOOFRE versão 2: ver e tratar 600 bilhões de palavras\*1

Etienne Brunet²; Laurent Vanni³, Cláudia Grijó Vilarouca⁴ (Tradução)

A empreitada de *Google Books* da qual nos demos conta na ocasião das JADT<sup>5</sup> de 2012 teve um impacto de um abalo sísmico planetário. Apropriar-se, com a ajuda de scanners, dos livros do mundo inteiro, não podia não motivar uma reação de defesa econômica, cultural e quase religiosa. E a cada um, cabia a defesa de suas águas territoriais contra as malhas do invasor. Projetos nacionais ou europeus foram construídos para aceitar o desafio, sem impedir o crescimento da Torre de Babel (que alguns chamam de Torre de *Babil*). Em dois anos, a altura dobrou,

-

NOTA DOS EDITORES: Revisão do original em francês: Carlos Maciel (UNIVERSITE NICE SOPHIA ANTIPOLIS); Deise J. Tarouco Freitas (UNIVERSIDADE FEDERAL DE SANTA CATARINA)

<sup>&</sup>lt;sup>1</sup> A versão original em francês deste capítulo foi publicada originalmente como artigo na revista *Texto Digital*, n.2, 2014.

NOTA DOS EDITORES: Revisão do original em francês: Carlos Maciel (UNIVERSITE NICE SOPHIA ANTIPOLIS); Deise J. Tarouco Freitas (UNIVERSIDADE FEDERAL DE SANTA CATARINA)

<sup>&</sup>lt;sup>2</sup> Professor da *Université Nice Sophia Antipolis*, fundador do *Laboratoire Bases*, *Corpus*, *Langage* (BCL/UNS), criador do *software Hyperbase*. CONTATO: brunet@unice.fr

<sup>&</sup>lt;sup>3</sup> Engenheiro do *Laboratoire Bases, Corpus, Langage* da *Université Nice Sophia Antipolis* (BCL/UNS). CONTATO: laurent.vanni@unice.fr

<sup>&</sup>lt;sup>4</sup> Pesquisadora do Núcleo de Pesquisa em Informática, Literatura e Linguística da Universidade Federal de Santa Catarina (NuPILL/UFSC). Contato: claudia.vilarouca@gmail.com

<sup>&</sup>lt;sup>5</sup> Nota da Tradutora: JORNADAS INTERNACIONAIS DE ANÁLISE ESTATÍSTICA DE DADOS TEXTUAIS.

tanto para o francês quanto para outras línguas<sup>6</sup>. Em si, essa dimensão tornada considerável poderia não mudar os perfis nem tornar caducas as análises anteriores. Mas, por um lado, o inchaço dos dados não foi homogêneo: certos períodos primitivamente quase vazios ganharam volume e o desequilíbrio entre os primeiros e os últimos períodos foi parcialmente corrigido<sup>7</sup>. Por outro lado, os textos já disponíveis na versão de 2009 foram retomados na base e submetidos a tratamentos melhorados ou radicalmente novos, embora o corpus de 2012 não seja sobrepujável ao precedente, nem em sua composição nem em sua exploração. E a análise desse projeto gigantesco deve ser retomada, sem a *priori*.

Para aqueles que abordam esta questão pela primeira vez, convém explicar que os milhões de livros espoliados são localizáveis e consultáveis na rede, por meio da grade do *Google Books* que filtra mais ou menos os dados, conforme as restrições do *copyright*. Ora, esse procedimento clássico, cuja entrada é uma palavra e a saída um contexto, pode ser acompanhado por um exame que se concentra sobre as frequências. A entrada ainda é uma palavra (ou um conjunto de palavras), mas a saída fornece um gráfico que reproduz a evolução das frequências da palavra considerada entre duas datas. Um sítio particular, *Culturomics*, é dedicado a essa exploração estatística do *corpus*<sup>8</sup>. Consultemo-lo a partir do par de substantivos *homem-mulher* que encabeçam frequentemente as listas de frequência.

<sup>&</sup>lt;sup>6</sup> Estamos em 89 bilhões para o francês, 349 para o inglês (cujas numerosas variedades/VARIANTES? podem ser isoladas), 53 para o alemão, 67 para o espanhol e 33 para o italiano, recém chegado. Esses números correspondem aos dados que podem ser baixados. Eles são superiores, como se pode ver no quadro 1 do artigo publicado em *Proceedings fo the 50th Annual Meting of the Association for Computational Linguistics*, p. 170. Três outros *corpora* estão disponíveis, dos quais não diremos nada devido à falta de conhecimento e de teclado: russo, chinês e hebraico.

<sup>&</sup>lt;sup>7</sup> Nota da Tradutora: A correção é mais clara para o francês do que para as outras línguas.

<sup>&</sup>lt;sup>8</sup> O paralelismo entre os dados de *Google Books* e os *corpora* de *Culturomics* diverge ao longo do tempo. Os primeiros são evolutivos e se enriquecem a cada dia. Os segundos são estáticos e representam um estado fixo dos primeiros, em uma determinada data.

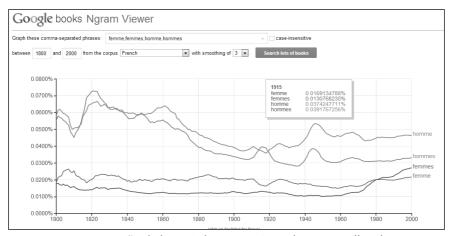


Figura 1. Interrogação da base *Culturomics*. O par *homem/mulher* de 1800 a 2000<sup>9</sup>.

Pode-se parar com o cursor em qualquer ano para interpretar os andamentos da história e constatar, por exemplo, o crescimento de *homens* em tempos de guerra.

### Melhorias e lematização

Poder-se-ia acreditar que o exame de Culturomics não variou desde a primeira versão e que somente a base de dados se estendeu. Na verdade, existe um modo de uso dito avançado<sup>10</sup> que muda radicalmente o ponto de vista e permite acessar não somente as formas

Assim, a versão 2012 é uma atualização da versão 2009. Outras atualizações estão previstas para os próximos anos.

<sup>&</sup>lt;sup>9</sup> Uma vez completadas as zonas de preenchimento (palavras procuradas, datas de início de fim, o *corpus* escolhido e a opção de correção), os dados disponíveis na rede encontram-se neste link:

https://books.google.com/ngrams/interactive\_chart?content=femme%2Cfemmes%2Chomme%2Chommes&year\_start=1800&year\_end=2000&corpus=19&smoothing=3

<sup>&</sup>lt;sup>10</sup> As múltiplas maneiras de consultar a base Culturomics estão detalhadas no endereço: http://books.google.com/ngrams/info

brutas, mas formas anotadas/inventariadas e providas de um código gramatical.

1 - Contrariamente ao corpus da edição 2009, os dados da edição 2012 se beneficiaram da lematização. Esta não está completa e não permite obter logo todas as formas, por exemplo, do verbo andar. Mas se o reagrupamento das formas não foi previsto, a separação dos homógrafos foi implementada e pode-se obter duas séries separadas para a forma "le" (artigo o) e forma "le" (o) pronome pessoal. Basta acrescentar a categoria à palavra buscada, por exemplo le\_DET ou le\_PRON. Uma dezena de sufixos está disponível para designar os códigos gramaticas. Grosso modo, eles são imutáveis de uma língua à outra. Note-se que esses códigos podem ser submetidos à avaliação, ou sozinhos ou em combinação com outros códigos ou palavras individuais. Pode-se assim, estabelecer a curva recapitulativa de todos reunidos construção os verbos ( VERB ) ΟU da preposição+determinante+substantivo (\_ADP\_\_DET\_\_NOUN\_) ou de todos os substantivos qualificados como "velhos" (vieux\_ADJ \_NOUN\_ + \_NOUN\_vieux\_ADJ). É admirável a potência do programa na figura 2, que reúne mais de 300 bilhões de palavras do corpus inglês e retoma as partes do discurso no tempo, de 1800 a 2000. Nunca, até aqui, a progressão do substantivo e do adjetivo e o declínio variável de todas as outras categorias tinham sido observadas em tal escala<sup>11</sup>.

<sup>&</sup>lt;sup>11</sup> A inclinação parece pouca e quase imperceptível para algumas categorias porque as variáveis são traduzidas em porcentagens. Porém, medidas em termos probabilísticos, as variações são bem consideráveis e muito significativas.

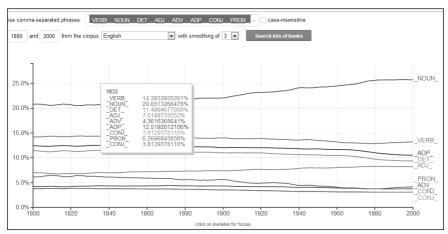


Figura 2. A evolução das partes do discurso no corpus inglês.

o2 – Aos códigos gramaticais se juntam alguns símbolos que definem a posição das palavras na frase (\_START\_ ou \_END\_) ou a relação das palavras entre elas ( => , <= , \_ROOT\_). Naturalmente, a avaliação pode recair não apenas sobre uma única palavra, lematizada ou não, mas também sobre uma cadeia de várias palavras (de 1 a 5), sendo que cada uma dentre elas pode admitir filtros. Em certos casos, no entanto, a menção acrescida de códigos gramaticais reduz o alcance a somente 3 palavras.

### 3 – Certas manipulações numéricas são possíveis para estipular:

- o agrupamento de várias palavras em somente uma solicitação (signe+);
- a relação proporcional de uma palavra a outra pela subtração (signe-) ou o quociente (signe / );
- ou o reequilíbrio/estabilização de duas palavras (signe \*, aplicado à palavra menos frequente).

Esses signos da metalinguagem são ativados, se for necessário, pelos parênteses (), ou desativados por chaves [].

4 – Os avanços não residem apenas na anotação dos textos nem na sofisticação da exploração. Eles vêm também das etapas iniciais que foram retomadas a partir das imagens escaneadas. Uma leitura óptica aperfeiçoada permitiu reconhecer os « s » longos das edições antigas ou, ao menos, corrigir por algum meio os erros sistemáticos que essa grafia gerava. Do mesmo modo, a segmentação (ou tokenização), que primitivamente se apoiava apenas nos espaços em branco, introduziu um separador no fim das frases, impedindo os *ngrams* de transgredir essa barreira. Ao contrário, a fronteira da página foi abolida. É preciso reconhecer tais mudanças, lamentando que as decisões sensatas não tenham sido tomadas logo de início.

#### Acesso mais leve e mais potente

Uma desvantagem, no entanto, impedia ainda o livre desenvolvimento da base *Culturomics*: ao obter uma curva opaca em lugar de uma série de quantidades, deparávamos com um terminus ad quem, que interrompia a cadeia dos tratamentos ulteriores. Podíamos nos sentir tão frustrados que percebíamos implicitamente os dados esperados: percorrendo com o mouse algum ano particular, víamos as porcentagens precisas surgirem em uma janela efêmera<sup>12</sup>, sem que pudéssemos retê-las e registrá-las.

1 – Os autores de *Culturomics* propuseram, então, uma API (*Application Programming Interface*) cujo *download* é possível, <sup>13</sup> que, por uma

<sup>&</sup>lt;sup>12</sup> A figura 1 mostra um zoom parcial realizado no ano de 1915 e a figura 2, no ano de 1833.

Todavia, não se trata de uma API verdadeira e certificada, mas da captação da mensagem retornada por *Culturomics* em resposta à solicitação. Os elementos que servem para estabelecer as coordenadas dos pontos da curva figuram nessa resposta e são apreendidos na transição. Esse desvio fica frágil e subordinado à estabilidade do diálogo servidor-cliente. Tal mudança interveio recentemente e tornou inoperante o primeiro programa (GetNgrams.py) distribuído pelos autores de *Culturomics*. Laurent Vanni, do laboratório BCL, se encarregou das retificações necessárias e poderá assegurar a manutenção desse serviço.

determinada palavra, distribui as 201 porcentagens observadas ao longo da cronologia. Esse pequeno programa, escrito em linguagem Python, pode ser facilmente modificado<sup>14</sup> e integrado em uma cadeia de tratamento. A janela abaixo mostra o resultado obtido quando o programa é lançado com seus parâmetros explícitos ou, na falta de, apenas um é exigido: a palavra buscada (no nosso caso *day*).

```
Microsoft Windows XP [version 6.1.7601]
(C) Copyright 1985-2001 Microsoft Corp.

C:\text{C:HYPERBAS}\text{python getngrams.py day}

http://books.google.com/ngrams/graph?content=day&year_start=1800&year_end=2000&corpus=19&smoothing=0&share=
http://books.google.com/ngrams/graph?content=day&year_start=1800&year_end=2000&corpus=19&smoothing=0&share=
year.day
1800.5.56453642275e-07
1801.6.52347694086e-07
1802.3.83127144232e-07
1803.1.2873090684e-07
1804.2.54569727076e-07
1804.2.54569727076e-07
1806.7.05102706088e-07
1807.4.230918762e-07
1807.4.230918762e-07
1809.2.559853201e-07
1810.3.01846057482e-07
1811.2.3589031728e-07
1811.2.3589031728e-07
1812.4.94908704241e-07
1813.2.81399337609e-07
1814.3.38726529662e-07
```

Figura 3. Resultado da API getngrams.py

Na realidade, a resposta de *Google Books* é uma lista de 201 porcentagens, cada uma representando um ano de 1800 a 2000 no corpus considerado. Para transformar essa frequência relativa em frequência absoluta e permitir assim o cálculo da variável reduzida, fizemos intervir a extensão de cada ano nesse corpus (cujos dados foram retirados em arquivos retroativos de *Culturomics*). Assim, sabendo que a extensão do ano 2000 é de 1.182.754.941 palavras no corpus francês de 2012, recebemos da API a distribuição por ano de uma palavra proposta, por exemplo, a palavra "amor" a partir da qual constatamos a proporção 0,00013133 para o mesmo ano 2000. Logo, a frequência real é de 1182754941 \* 0.00013133 = 155331 para a palavra "amor" no ano 2000. Na verdade, por razões de legibilidade

\_

<sup>&</sup>lt;sup>14</sup> Um pequeno retoque, graças a Laurent Vanni também, converte na codificação ANSI tradicional os caracteres acentuados do unicode.

renunciamos aos mínimos detalhes dos 201 anos, reunindo-os por partes equilibradas. Essa opção aperfeiçoa melhor a partição, deslocando as marcas cronológicas de modo a igualar ou, ao menos, harmonizar o peso de cada parte. Com efeito, acontece que, na composição do *corpus* os livros das épocas antigas são menos representados que os livros modernos porque são menos numerosos e menos disponíveis nas bibliotecas.

Retomando o exemplo do par homem/mulher representado na figura 1, dispomos assim de elementos dos quais se alimenta habitualmente a lexicometria (frequências absolutas distribuídas em um quadro com linhas correspondentes às palavras e colunas reservadas aos textos ou aos períodos). Por conseguinte, todas as transformações são possíveis, levando aos histogramas (de uma linha ou de uma coluna) e às análises fatoriais ou em árvores.

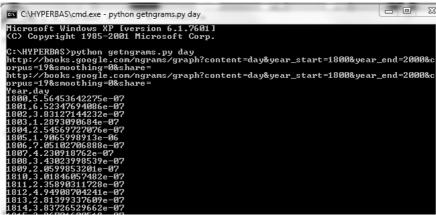


Figura 4. A constituição de um quadro de frequências reais (ou absolutas)

A constituição de tais quadros não é das mais rápidas, pois cada linha resulta de uma consulta específica, lançada na rede. Deve-se prever um ou dois segundos para cara uma. No entanto, não há nenhuma limitação nos parâmetros de busca, que pode misturar, sem problema, as formas brutas, as formas codificadas, as categorias consideradas em

seu conjunto, as restrições no ambiente prosódico, gramatical ou semântico. Podemos reunir no mesmo conjunto palavras simples (ou *unigrams*) ou expressões complexas (de 2, 3, 4 ou 5 palavras). Podemos introduzir no mesmo quadro observações provenientes de corpora diferentes com a condição de que a periodicidade continue constante e que as marcas cronológicas não sejam deslocadas<sup>15</sup>.

2- Assim, aos quatro representantes franceses do quadro 4, juntemos seus homólogos ingleses, espanhóis e italianos, reunindo num mesmo total as mulheres de todas as nacionalidades e, simetricamente, o conjunto internacional dos homens. A evolução que constatamos é definitiva. Se a massa dos homens é o dobro da massa de mulheres (221 milhões de ocorrências contra 548 milhões), sua superioridade é reduzida ao longo dos anos. As mulheres, sete vezes menos numerosas no início, têm a desvantagem diminuída, ao ponto de ultrapassar os homens na última parte (aproximadamente 10 milhões contra 8 milhões de homens). Se sobrepusermos as duas curvas como na figura 5, a inversão das tendências é manifesta

\_

<sup>&</sup>lt;sup>15</sup> A lei hipergeométrica normalmente utilizada em lexicometria é inutilizável na escala do bilhão. Por essa razão, recorremos à lei normal e ao cálculo clássico de desvio reduzido. Note-se que as curvas e as análises fatoriais ou em árvore se apoiam nesses desvios, que sempre se referem à totalidade do *corpus* considerado. Não iremos considerar o quadro de frequências como um quadro de contingência, que se bastaria e cujos totais secundários permitiriam estabelecer os efetivos teóricos e as medidas do CHl2. Tal cálculo poderia ser legitimado (temos facilidade para a análise fatorial), porém, preferimos, a fim de se ter mais abrangência e estabilidade, considerar sempre o *corpus* inteiro como a referência interna pelas partições cronológicas, das quais ele é o total. No caso de *corpora* diferentes tratados ao mesmo tempo, essa ponderação é indispensável, pois, do contrário, os resultados seriam somente o reflexo do tamanho dos *corpora* e *subcorpora*.

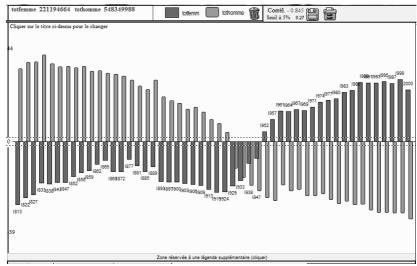


Figura 5. Homens e mulheres em cinco línguas ocidentais. Evolução inversa.

Mantendo as 20 linhas e as 50 colunas do quadro, podemos recorrer à análise em árvore. Ao se unir a tantas publicações que dizem respeito aos *Estudos de Gênero*, a figura 6 oferece uma confirmação interessante, já que ela trata da oposição *homem/mulher* em diferentes línguas<sup>16</sup>. Dessa forma, temos a rara ocasião de cruzar o tempo e o espaço e observar se as fronteiras nacionais impedem um movimento de propagar suas ondas no mundo inteiro.

Podíamos pensar que as línguas imporiam suas fronteiras. *Certamente*, há uma solidariedade nacional entre os homens, o singular e o plural que se reconhecem como compatriotas e que se apertam as mãos. Todavia, este é apenas um critério secundário. Sabendo-se que, de modo geral, em dois séculos, o plural tende a se apagar diante do singular (é o que observamos nos determinantes, sobretudo na relação

\_

<sup>&</sup>lt;sup>16</sup> Pode-se fazer que uma mesma palavra, por exemplo, um substantivo próprio, seja comum a vários *corpora*. Ver-se-á então a popularidade do personagem ou do topônimo evoluir tanto na geografia como na história. Em tais casos um símbolo inicial é útil para a distinção das línguas (*F*= french, *E*= english, *G*= german, *S* = spanish, *I* = italian).

de le+lα+l' versus les [o+a+o/a versus os/as]), podíamos imaginar que tal distinção da quantidade se encontraria em palavras tão correntes. Não é bem assim. Nem a quantidade nem a língua impõem sua dominação. É o gênero ou, mais precisamente, o sexo que é decisivo e que põe em confronto dois campos irreconciliáveis: os homens formam bloco na parte de baixo do gráfico; as mulheres, acima. Temos a situação radical de Sodoma e Gomorra, mesmo se uma palavra se aproxima sem receio da linha de demarcação: a mulher francesa que está de olho na direção dos homens. Isso se deve talvez à ambiguidade da palavra mulher que reúne dois papéis que o inglês distingue com woman e wife. Enquanto wife, a mulher francesa tem menos razão para se opôr à classe masculina. É a ocasião de destacar que a quantidade não dissipa as incertezas. Mesmo em um caso aparentemente simples, é perigoso ultrapassar a ponte entre as duas línguas: as equivalências que estabelecemos entre termos simétricos brutalizam sempre de algum modo a realidade semântica.

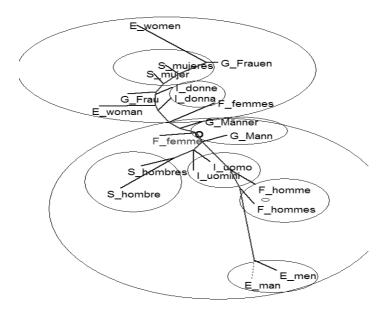


Figura 6. O par *homem/mulher* em cinco línguas ocidentais Análise em árvores baseada em 20 palavras e 762 milhões de ocorrências.

Diante de tal quadro, a análise fatorial (figura 7) propõe alguns complementos. Esse método possui a vantagem de juntar linhas e colunas e de interpretar umas com o auxílio das outras. Quanto à coluna, as coisas são claras: a cronologia reina absolutamente. Todas as partes de 1800 a 1933 estão à direita; todas as outras, à esquerda. Quanto às linhas, como na análise precedente, as escolhas são ordenadas e os campos reordenados, as mulheres à esquerda, os homens à direita. Aqui também a mulher francesa joga com o limite e confirma sua propensão à independência, exceto à traição. O primeiro fator, que ocupa 66% da variância, coloca os homens em partes distanciadas e as mulheres em partes recentes e nos dois casos, ele distingue o singular e o plural, sendo que este último ocupa a posição extrema. A oposição parece então menos forte entre o homem e a mulher do que entre os homens e as mulheres. Os debates acerca da posição/lugar do homem e da mulher na sociedade seguramente utilizar o valor generalizante do singular, porém, frequentemente a constatação das desigualdades se faz à luz das classes coletivas: os operários, os trabalhadores, os ricos, as mulheres.

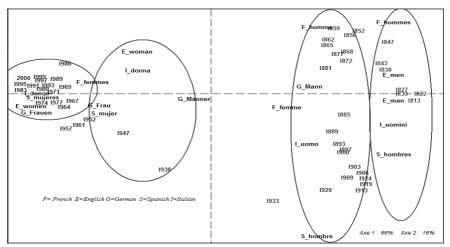


Figura 7. Análise fatorial do mesmo par homem/mulher

3 – O exame cruzado de *corpora* diferentes é mais sólido quando, escapando das aproximações da tradução, se propõe a análise dos elementos estáveis que não se alteram de uma língua para outra. É o caso das pontuações, dos topônimos, dos nomes de pessoas e, em um grau menor, das partes do discurso<sup>17</sup>. Observamos tanto a convergência, tanto os particularismos nacionais. O exemplo dos sinais de pontuação é típico do primeiro caso, de um movimento uniforme, semelhante à deriva dos continentes. Por toda a parte a obsolescência toca o ponto e vírgula, que figura, no entanto, entre os mais antigos signos do sistema. Mesmo declínio dos signos no qual a expressão se faz mais intensa e mais direta (exclamação e interrogação<sup>18</sup>). Estes sinais aproximam-se das etapas mais distantes no tempo, que se encontram na figura 8, à direita. Já o ponto, os dois pontos e as reticências encontram-se à esquerda, seguindo assim a tendência que manifestam as etapas cronológicas mais recentes.

A que podemos atribuir essa espécie de glaciação universal que se contenta com uma expressão mais neutra, mais fria, mais voltada para a constatação que para a emoção? Sem dúvida, menos à evolução das línguas que a uma mudança na composição dos corpora. As publicações mais recentes, que são também as mais numerosas, não foram submetidas à triagem da história: é a irrupção da edição, na qual pululam obras de informação, tratados técnicos e temas dos mais diversos. Os livros mais antigos sobreviveram ao esquecimento e à perdição porque, mantido o interesse por eles, apareceram reedições que aumentaram a chance deles de sobrevida. Eis o privilégio das obras

\_

<sup>&</sup>lt;sup>17</sup> O mesmo jogo de oito etiquetas gramaticais serviu para todos os *corpora*: NOUN, VERB, PRON, ADJ, ADV, DET, CONJ e ADP (=preposição + posposição). Não é certo o fato de que convenha de maneira homóloga para todas as línguas. Juntam-se dois códigos de menor interesse: PRT (siglas) e X (palavras estrangeiras).

<sup>&</sup>lt;sup>18</sup> A figura 8 não faz menção do ponto de interrogação, afastado por prudência, devido à especificidade de seu emprego em espanhol. Na verdade, a distorção não se produz e o ponto de interrogação é classificado sem rodeios ao lado do ponto de exclamação. Em compensação, a vírgula não pôde ser considerada na pesquisa: como esse signo pertence à metalinguagem do motor de busca, ele escapa a toda investigação, o que vale igualmente para as aspas.

literárias; raramente é o caso das publicações técnicas, que o progresso condena rapidamente.

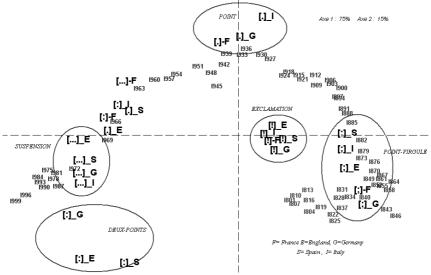


Figura 8. Análise fatorial dos sinais de pontuação

Tanto a ortografia dos topônimos como a dos personagens históricos varia pouco ao passar as fronteiras. Com alguns ajustes<sup>19</sup>, podemos medir a notoriedade variável que um nome próprio de lugar ou de pessoa pode ter em um determinado país e em certa época.

<sup>&</sup>lt;sup>19</sup> Rome se diz Roma em italiano e London, Londres em francês. Mas nem Paris, nem Berlin, nem Madrid admitem variantes nacionais.

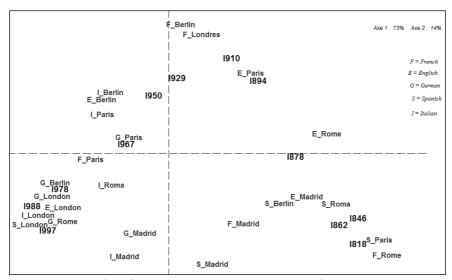


Figura 9. Análise fatorial de cinco capitais em cinco línguas europeias.

Tomemos o exemplo de cinco capitais europeias que corresponde às cinco línguas disponíveis. Os mesmos nomes, nas mesmas divisões do tempo, são sucessivamente buscados nos cinco corpora estabelecidos em 2012. A análise fatorial (figura 9) destaca um deslocamento de interesse histórico que tende a abandonar as cidades do sul (Roma, Madri e Paris) para se unir às cidades do norte (Berlim e Londres). Contudo, o chauvinismo tende a contrariar essa tendência quando a capital do país está em questão: mesmo nas divisões recentes, *Paris* continua popular na França, e *Roma* na Itália (os pontos I\_Roma e F\_Paris estão à esquerda do gráfico, como E\_London e G\_Berlin, estão onde se concentram as divisões contemporâneas).

A tentação é grande em avançar e em não se contentar com sondagens, mesmo através de palavras ou signos muito recorrentes. Pensemos no conjunto das palavras e o conjunto dos *corpora* ocidentais, a saber, aproximadamente 600 bilhões de observações elementares. O objeto de estudo retoma e desenvolve o estudo das partes do discurso, já abordada na figura 2, consagrada ao *corpus* inglês. Ora, esse *corpus* não

tem nada de específico: as mesmas tendências e os mesmos acidentes se encontram em outras línguas; uma tendência clara que se volta a favor das categorias nominais (nomes e adjetivos) em detrimento do verbo e seus acólitos (advérbios, pronomes, conjunções). Nesse caso, o julgamento é ainda mais perplexo: trata-se de um movimento de fundo, de natureza propriamente linguística? Não seria mais um simples artefato previsível, uma consequência da lei dos gêneros? Se a coleção de dados se fez ao acaso das oportunidades, não corremos o risco de encontrar o utilitário e o diverso na proximidade imediata, e objetos, mais raros, de um valor superior, na herança antiga? Teríamos então a oposição bem conhecida entre o literário e o utilitário, entre a expressão e a informação, que se traduz por uma combinação diferente do verbo e do nome.

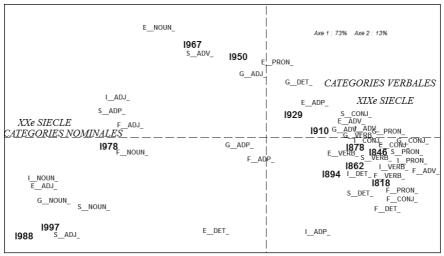


Figura 10. Análise fatorial das partes do discurso (cinco corpora, 600 bilhões de palavras)

Parece, no entanto, que fenômenos de simplificação estão na obra tanto na sintaxe quanto na pontuação. Não apenas a frase se destitui de suas construções carregadas que giram em torno do verbo (conjunções e pronomes relativos, principalmente), mas também a linguagem parece fazer economia do pequeno aparato que usualmente acompanha o nome (determinantes e preposições). Medidas complementares parecem prová-lo quando duas construções são postas lado a lado, por exemplo, a construção *Prép + Nom*, preferida a *Prép+Dét+Nom*; ou *Verb+Nom* melhor que *Verb+Dét+Nom*. Se as preposições hesitam entre a subida (em espanhol), a estabilidade (em alemão) e a descida (por toda a parte fora), o decréscimo dos determinantes é geral, enquanto que por toda a parte os substantivos e os adjetivos são em progressão. Nunca, até o momento, tinha sido observada essa disjunção do grupo nominal, com o substantivo sozinho à frente, seguido à distância do adjetivo e, mais distante, preposições e determinantes que encontram dificuldade para seguir a progressão.

## Dúvidas e vertigens

Todavia, do alto da Torre de Babel, com 600 bilhões de palavras amontoadas sob os pés, eis que surgem a vertigem e a dúvida. As nuvens que envolvem o cume escondem as fundações. Como se assegurar que o edifício irá resistir à maldição, à discórdia linguística que se abateu sobre a Torre da Bíblia?

Os autores de *Culturomics* jogaram honestamente o jogo da transparência, permitindo a cada um verificar a exatidão dos números. Não se trata somente de controlar os retornos ao texto: na falta de tempo ou de paciência uma sondagem dessa ordem apenas pode se efetuar em um filamento infinitesimal dos dados, e se a referência anunciada da aventura se revelava falsa ou ausente, nenhuma conclusão geral poderia ser retirada desse defeito particular. Um verdadeiro controle somente pode ser realizado retomando os cálculos na base, o que supõe que se tenha acesso aos dados brutos. Claro que podemos restaurar o processo até as imagens escaneadas e é preciso se contentar com os extratos e dados numéricos obtidos que foram conseguidos – e que podem ser baixados. Não é necessário se assustar com a quantidade e o tamanho dos arquivos a baixar, se nos satisfazemos com os *unigrams* (ou palavras individuais). Aqui,

felicitamo-nos por um progresso notável com relação aos dados de 2009: os arquivos são classificados segundo a inicial das palavras. Resta, para cada letra, proceder às operações carregadas de triagem, de compactação e de reagrupamento em uma única e mesma base. Os autores do *Culturomics* até mesmo recomendam essa duplicação, com o objetivo de desenvolver a exploração de seus dados aliviando o servidor deles.

concentrando os dados, neutralizando distinção Mesmo maiúsculas/minúsculas, reduzindo 200 anos divisões 12 cronológicas, afastando as palavras raras que têm grande chance de ser erros de leitura<sup>20</sup>, chegamos, apenas com números e sem nenhum texto, a uma base de 300 milhões de octetos, com consideráveis 1,5 milhões de entradas. Obviamente que o objetivo dessa operação custosa não é apenas controlar os dados, mas, sobretudo, permitir uma exploração fácil e imediata, sem os pesos e lentidões associados à rede. Tivemos especialmente a preocupação de nos livrarmos das porcentagens simples e de encontrar os números ou efetivos absolutos, que permitem o desdobramento pleno da estatística e da aplicação dos métodos multidimensionais. Veremos, mais além, uma breve exposição dessa base, quando o exame dos dados já tiver sido feito.

Uma discordância inicial já nos inquieta: os números que notamos nos dados baixados não correspondem exatamente àqueles que obtemos pela rede. Retomemos o exemplo da palavra homem, cujo total chegaria a 43 milhões (precisamente 43.424.969) no quadro 4, extraído de *Culturomics*. Contamos quase dois milhões a mais (exatamente 45.191.302) nos dados transferidos. Disso deve-se concluir que as contas definitivas fixadas em 2012 foram sujeitas a correções e que o tamanho de cada ano do *corpus* foi calculado após a eliminação dos refugos. Certamente, o perfil cronológico da palavra não sofreu grandes

<sup>&</sup>lt;sup>20</sup> Colocamos a linha em 100 ocorrências, muitas acima daquela do *Culturomics*, que é de 30. O tamanho do *corpus* foi reduzido a 70 bilhões de palavras na área de francês.

perturbações, mas a aproximação diminui um pouco o crédito que damos aos extratos obtidos.

Quanto ao volume dos refugos, não podemos dar uma estimativa, pois eles foram suprimidos numa depuração prévia aos resumos disponíveis. Porém, teremos uma ideia a partir de um trecho da figura 11 que relaciona todas as variedades encontradas quando o exame concerne à palavra été. Com dois acentos no espaço de três letras, os avatares ortográficos se multiplicam ao infinito e uma centena de atrofias lexicais nasceu dessa proliferação desordenada.

249614	79049	99	276
ete	ete_ver	etè_ad	éte_pro
130 ete_	b	<b>v</b>	n
5601	3271 ete_x	119	37184
ete_adj	126767 eté	etè_no	éte_ver
95	123 eté_	un	b
ete_ad	111	91	58 éte_x
р	eté_ad	etè_ver	129257681
41892	р	b	été
ete_ad	20076	75 etê	444 été_
V	eté_ad	26 etê_adv	410
62	V	66 etë	étéve
ete_con	48605	64022 éte	rb
j	eté_no	3960	2171786
16211	un	éte_adj	été_no
ete_det	1493	12181	un
99993	eté_pro	éte_ad	127085852
ete_no	n	V	été_ver
un	54731	8065	b
2853	eté_ver	éte_det	962 étè
ete_pro	b	2103	53 étè_det
n	1485 eté_x	éte_no	
78 ete_prt	415 etè	un	

41	131	297	8954
étè_no	ète_det	ête_ad	êté_ver
un	1310	V	b
823	ète_no	769	42 êtè
étè_ver	un	ête_det	120 êtê
b	131	1511	39 êtê_adv
345 étê	ète_pro	ête_no	39
48 étê_det	n	un	êtê_ver
39	366	148	b
étê_no	ète_ver	ête_pro	234 ëte
un	b	n	46
212	40 ète_x	41 ête_prt	ëte_ad
étê_ver	981 èté	3224	V
b	51	ête_ver	75 ëte_det
773 étë	èté_no	b	71
81 étë_adv	un	66 ête_x	ëte_no
109	923	9955 êté	un
étë_det	èté_ver	114	614 ëté
36	b	êté_adj	39 ëté_det
étë_no	1000 ètè	354	531
un	67 ètè_adv	êté_ad	ëté_ver
515	135	V	b
étë_ver	ètè_det	102	109 ëtë
b	75 <b>1</b>	êté_det	44 ëtë_adv
2482 ète	ètè_ver	355	
459	b	êté_no	
ète_ad	6259 ête	un	
V	90 ête_adj		

Figura 11. Os avatares da palavra été

O leitor óptico pode ser responsável por um terço desses erros: cada uma das duas vogais da palavra, podendo admitir seis interpretações, há então 36 combinações possíveis dentre as quais nenhuma foi negligenciada<sup>21</sup>. Os dois terços restantes vêm da lematização. Há, primeiramente, uma repetição mecânica, que explica o segundo terço. Cada forma deve poder ser examinada sob duas formas: com e sem codificação gramatical. Resta o último terço imputável às más escolhas do parser\*. Em princípio, há somente duas opções possíveis para été. Ou nos referimos à estação estival (verão) ou é o verbo ser no particípio passado. Ora, é exatamente o que constatamos quando a grafia é correta: sem menção de código, a palavra possui 129.257.681 empregos, dentre os quais 2.171.786 aparecem como nomes e 127.085.852 como verbos. Falha apenas uma guantidade insignificante: 43 ocorrências. Parece que há como apaziquar o linguista. Entretanto, essa exatidão providencial pode incomodá-lo, ao observar o pânico do lematizador confrontado às formas desconhecidas e distribuindo códigos ao acaso. No quadro 11, encontramos o jogo completo dos códigos disponíveis, sem ter em conta os das conjunções. Como conciliar dois comportamentos tão diferentes da máquina: um desarranjo irremediável nas situações confusas e um corte brusco em contextos que são tidos como claros? O primeiro é fácil de compreender: quando uma palavra é mal compreendida ou mal interpretada, o lematizador perde suas referências e cai num erro que, por sua vez, gera um segundo. E por receio de que o leitor óptico se engane novamente, entramos num labirinto inextricável no qual o fio se perde. O segundo tratamento se deve provavelmente a decisões autoritárias que obrigam a escolha na alternativa. É preciso dar crédito à boa fé do primeiro tratamento que se engana frequentemente, mas de maneira aleatória, ou à autoridade peremptória do segundo que pode conduzir ao erro sistêmico?

\_

<sup>\*</sup> programa de computador

Devemos supor correta a leitura de três letras, independentemente dos acentos. Na verdade, muitos dos *e* reconhecidos são falsos, são *a* ou *o* disfarçados. E, ao contrário, muitos verdadeiros *été* foram perdidos no labirinto do alfabeto.

Para decidi-lo, vamos até o fim do alfabeto, lá onde os lexicógrafos, já terminando o seu trabalho, ficam menos atentos. Observemos a palavra ver, com a ortografia correta. A análise não parece ter sido supervisionada já que encontramos vários códigos fantasiosos que a gramática francesa não permite, não mais que a espanhola que dá outro sentido à mesma palavra.

758187 ver 42188 ver\_adj 135 ver\_adp 526 ver\_adv 9446 ver\_det 528603 ver\_noun 1503 ver\_pron 3506 ver\_prt 88034 ver\_verb 83771 ver\_x

Em compensação, a palavra vers [prep. "na direção de"; verso; verme] recebeu um tratamento diligente que dá todas as ocorrências (há 38 milhões) como preposição, esquecendo os poetas que fazem versos e os mortos que fazem vermes. Percebemos aí a prova de um tratamento duplo. O singular ver entregou-se inocentemente às máquinas, que o amputaram, deixando de lado certas formas e ocorrências, e se encontrou bizarramente fragmentado, com um quociente de boas respostas de 2 para 3 (528.603 para 758.187), apesar de tudo. O plural ficou sujeito a uma decisão automática, sem levar em consideração o contexto. Ora tais decisões arbitrárias tocam todas as palavras frequentes. Observamos todas as palavras cuja frequência ultrapassa o milhão. Ora se encontrou 6.001 que possuem um código único – o que é fácil de notar já que a palavra sem código e a palavra com código tem a mesma soma. É impossível que um lematizador justo não tenha notado - certo ou errado - homógrafos nessa porção enorme, que representa a grande maioria das ocorrências, a saber, 60 bilhões para 70.

### Esperanças e perseverança

Podíamos fazer de outro modo? Podemos duvidar, devido à dupla restrição de uma quantidade gigantesca para tratar em um tempo limitado. Os dados de *Frantext* possuem, seguramente, uma fiabilidade superior, graças a uma apreensão manual e a controles reiterados durante mais de quarenta anos. Antes de a informática ter sido solicitada para tratar os dados, equipes de linguistas tinham balizado o terreno e inventoriado as dificuldades e as soluções. E o objetivo inicial – a fabricação de um dicionário – era claramente linguístico. A empresa de *Google Books* é apenas documentária, mas sua ambição é sem limites. Enquanto o *Frantext* se contentava com milhões, *Culturomics* toca bilhões. Podemos esperar que um dia a qualidade acompanhe a quantidade.

Porém, em dois anos, quanto caminho percorrido para um e para outro objetivo. Vamos dar apenas um exemplo, tirado do verbo être (figura 12). Agora que as formas verbais foram desambiguizadas, embora grosseiramente, uma segunda basta para reunir no mesmo quadro todas as grafias que pertencem ao modelo, mesmo as formas homógrafas como été, être ou étais. Apesar da necessidade em todo discurso e de sua presença em cada frase, esse verbo cuja frequência chega a 1,4 bilhão está em declínio regular (correlação cronológica = -0,96), como a maioria dos verbos. Sobrevivem, no entanto, as formas impessoais do particípio e do infinito (étant, été e être) e aquelas do presente (sui, es, est, sommes, sont). Todo o resto é eliminado em épocas distantes e acompanha o século XIX. À simplificação da pontuação e da sintaxe, acrescenta-se a da conjugação . Porém, se a análise em árvores da figura 12 planta o verbo no tempo com a cabeça no presente e as raízes no passado, futuro e condicional ocupando o espaço intermediário, ela também é sensível às pessoas, sendo que a terceira possui as duas pontas da cadeia, enquanto que uma esfera compacta concentra as duas outras.

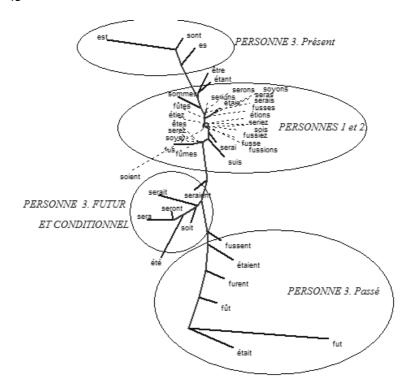


Figura 12. Análise em árvores das formas desambiguizadas do verbo *être* [ser]

Milhares de gráficos semelhantes podem ser extraídos da base GOOFRE2 que nós oferecemos ao público de forma inédita<sup>23</sup>. Que aproveitemos desde já. A esperança de vida de tais produtos é curta: uma nova versão de *Culturomics* está sendo preparada.

<sup>&</sup>lt;sup>23</sup> É tarde demais para detalhar a estrutura e explicar seu modo de uso. Expliquemos apenas que essa base permite três modos de exploração: 1- obtenção de curvas na Internet (como nas figuras 1 e 2); 2- obtenção de dados numéricos na rede com exploração local (figuras 3 a 10); 3- exploração inteiramente local do *corpus* francês (figura 12).

#### **REFERÊNCIAS**

Erez Lieberman, J.B. Michel, Joe Jackson, Tina Tang & Martin A. Nowak, Quantifying the evolutionary dynamics of language, *Nature*, Nature publishing Group, 2007, p. 713-716.

J.B. Michel & al., Quantitative Analysis of Culture Using Millions of Digitized Books, *Science*, déc. 2010, http://www.sciencemag.org/content/331/6014/176. full.html\*related.

Etienne Brunet, « Au fond du GOOFRE, un gisement de 44 milliards de mots », JADT 2012, pp. 7-21. *La base GOOFRE est téléchargeable à l'adresse*: http://logometrie.unice.fr/pages/bases et sur le site http://ancilla.unice.fr/GOOFRE.EXE.

J.P. Delahaye, Nicolas Gauvrit, Culturomics. *Le numérique et la culture*, Odile Jacob, 2013, 224 pages.

John Bohannon, *Google Books*, *and the future of Culturomics*, http://www.terceracultura.net/tc/wp-content/uploads/2011/01/culturomics.pdf John Bohannon, The Science Hall of Fame, The Gonzo scientist, *Science*, 14 jan. 2011, http://www.sciencemag.org/content/331/6014/143.3.full







