

# POR UMA FILOLOGIA VIRTUAL: O CASO DAS ATAS DA CÂMARA DE SÃO PAULO (1562-1596)

Vanessa Martins do MONTE  
Universidade de São Paulo (USP)

Maria Clara PAIXÃO DE SOUSA  
Universidade de São Paulo (USP)

## RESUMO

*Este artigo tem como objetivo discutir as reflexões teóricas e suas consequências práticas advindas do encontro de duas trajetórias de pesquisa: uma centrada na reflexão teórica sobre a natureza do texto digital e a contingência do trabalho filológico no meio digital. A outra no trabalho filológico centrado nas Atas da Câmara de São Paulo escritas entre 1562 e 1596. Busca-se amalgamar essas duas trajetórias no ponto que encontramos em comum: o objetivo de trabalhar o texto levando em conta sua época, sua língua, seus autores e leitores, e recriá-lo para o olhar de novos leitores, em uma nova época, sob uma nova luz. Sugerimos, neste artigo, o termo 'edição filológica virtual' para descrever o trabalho de edição iniciado com as Atas da Câmara de São Paulo, mostrado em detalhes neste artigo.*

## ABSTRACT

*This paper aims to discuss the theoretical reflections and their practical consequences that came from the convergence of two trajectories of research: one centered in the theoretical reflection about the nature of the digital text and the contingency of the philological labor in the digital environment. The other centered in the philological labor around the meeting minutes of São Paulo's village written between 1562 and 1596. One searches to combine these two trajectories in their shared point: the objective of working with the text considering its period, its language, its authors and readers, recreating it for the approach of new readers, in a new period, under a new light. We suggest the term 'virtual philological edition' to describe the work of edition started with the meeting minutes of São Paulo's village, laid down in details at this paper.*

## PALAVRAS-CHAVE

*Filologia. Edição virtual. Humanidades Digitais. Atas de São Paulo.*

## KEYWORDS

*Philology. Virtual Edition. Digital Humanities. Meeting Minutes of São Paulo.*

## Introdução

“**Virtual:** fisicamente inexistente, e sim criado por programas de computação, para parecer real aos sentidos (diz-se de imagem ou ambiente) <imagem n.> <realidade n.>” (HOUAISS, 2012).

Este artigo nasce do encontro de duas trajetórias de pesquisa: uma centrada na reflexão teórica sobre a natureza do texto digital e a contingência do trabalho filológico no meio digital (PAIXÃO DE SOUSA 2009, 2013a, 2013b) e a outra no trabalho filológico focado nas *Atas da Câmara de São Paulo* escritas entre 1562 e 1596 (MONTE & FACHIN a sair, MONTE 2015, 2016a, 2016b). Na trajetória teórica, temos refletido sobre os impactos das tecnologias digitais sobre o trabalho filológico explorando não apenas as perguntas que só podem ser *respondidas* pela filologia no meio digital, mas também as perguntas que só podem ser *formuladas* pela filologia neste meio. Na trajetória do trabalho filológico, temos investigado meios para a decifração, edição, publicação e difusão de um conjunto documental raro, visando superar os obstáculos trazidos pela natureza material e paleográfica dessa documentação e pela trajetória de sua transmissão desde o seu surgimento em meados do século XVI.

Um e outro olhar se cruzam na dupla pergunta proposta e explorada neste artigo: que possibilidades se abrem para o trabalho filológico sobre as *Atas da Câmara de São Paulo* (doravante ACSP) a partir das tecnologias digitais, e de que forma as possibilidades assim abertas contribuem

para a reflexão sobre a confluência entre as tecnologias digitais e o trabalho filológico? A exploração dessa dupla pergunta tocou nos pontos nevrálgicos do *conceito material de texto digital* esboçado em Paixão de Sousa (2013a): acreditamos que as perspectivas de decifração e difusão das ACSP em meio eletrônico só podem ser plenamente exploradas se compreendermos, como principal característica do texto em meio digital, sua natureza de *representação* artificial. O conceito material de texto digital aborda o texto segundo os aspectos das suas condições de produção, difusão e apreciação como objeto cultural, e traz a foco o processamento lógico artificial como seu ponto de diferenciação (dialogando, neste aspecto, com Eggert (2010)). Defende, ainda, que o texto digital em sua dimensão de ‘objeto humanamente legível’ é uma instância de ‘representação do conhecimento’ (nos termos de Unsworth 2001, 2004 e Clement, 2011). A natureza *fragmentária* do texto digital (trabalhada, principalmente, por Chaudiron, 2008) e os impactos dessa natureza sobre a edição filológica em meio digital são, a nosso ver, o ponto fulcral desse debate, com impactos importantes sobre a reflexão aqui proposta, como mostraremos.

A chave, nesta reflexão, está na ideia de *representação*: o ‘texto’ digital é uma simulação computacional de texto, destituída de realidade física, e plena de artifícios lógicos. É nesse sentido que a ‘edição filológica digital’ pode ser ressignificada como uma **edição virtual**, ou seja, como a apresentação *‘fisicamente inexistente’* e *‘criada por programas de computação para parecer real aos sentidos’* de um documento fisicamente existente no espaço e no tempo. No caso específico das ACSP, trabalha-se com um documento cuja existência física apresenta-se atualmente em tal estado de fragilidade que sua manipulação é contra-indicada. E que clama, portanto, por uma representação virtual. Do ponto de vista do trabalho filológico, a representação virtual do texto não é uma fraqueza: ao contrário, é sua maior vantagem com relação às outras formas conhecidas de texto. É este engenho na ilusão do olhar que torna a

representação computacional uma ferramenta poderosa de preservação da fidedignidade de um texto original. Discutiremos aqui como opera a conjunção aparentemente paradoxal entre a fidelidade ao texto almejada pelo fazer filológico e a simulação de texto almejada pelo processamento digital.

Assim, buscamos amalgamar as duas trajetórias motivadoras deste trabalho – a reflexão teórica sobre o digital e a dedicação ao fazer filológico – no ponto que encontramos em comum: o objetivo de trabalhar o texto levando em conta sua época, sua língua, seus autores e leitores, e recriá-lo para o olhar de novos leitores, em uma nova época, sob uma nova luz.

## **1. As Atas da Câmara de São Paulo (1562-1596) e seus desafios filológicos**

As ACSP, cujo primeiro exemplar conservado data de 1562, em conjunto com as atas da câmara de Santo André da Borda do Campo, lavradas entre 1555 e 1558, são documentos de inestimável importância histórica, cultural e linguística. Essa documentação, conservada atualmente no Arquivo Histórico Municipal Washington Luís, tinha inicialmente uma função estritamente documental: a de materializar textualmente uma função administrativa. Com o passar do tempo e o aumento da importância político-econômica de São Paulo, os livros de vereações vão adquirindo valor histórico.<sup>1</sup> Esse processo de valoração culmina com a publicação, em 1914, da edição impressa das atas, reeditada em 1967 com o acréscimo de um prefácio à segunda edição. Deve-se ressaltar que, a par da relação de parte das abreviaturas constantes dos originais, não há informação sobre as normas de transcrição utilizadas. Breve apresentação da história da transmissão das atas e da notícia de

<sup>1</sup> Segundo Karla Maestrini, no Simpósio Caminhos de Pesquisa no Arquivo Histórico de São Paulo: Uma Visão Multidisciplinar, realizado em 06/10/2016, o documento não nasce histórico, ele se torna histórico.

sua edição atual foi realizada em Monte & Fachin (a sair), texto com o qual dialogaremos neste artigo. Pode-se afirmar que, a partir da publicação da ‘versão’ impressa das atas, os originais são silenciados, ao ponto de as atas quinhentistas da câmara de São Paulo tornarem-se praticamente sinônimo de sua edição impressa de 1914/1967.<sup>2</sup> É compreensível que as edições impressas tenham sido as responsáveis por ampliar o acesso às atas, já que a documentação original oferece muitas dificuldades de decifração, tanto pela letra gótica cursiva, altamente encadeada, quanto pelo estado de conservação do suporte, muito fragmentário em vários fólios e com manchas e trechos corroídos. Colaborou para esse processo de utilização dessas edições como via de acesso exclusiva às atas quinhentistas o alto grau de confiança depositado no trabalho de decifração comandado por Washington Luís, Afonso d’Escragnolle Taunay e seus contemporâneos, ainda que não tenham anunciado seus critérios editoriais e o grau de fidedignidade mantido em relação aos originais, como apontam Monte & Fachin (a sair).

A ação do tempo associada às condições sabidamente pouco ideais de armazenamento causaram danos irreparáveis ao papel que serve de suporte material às ACSP. O papel, sem marca d’água até onde pudemos verificar, encontra-se extremamente fragmentado, com perdas de suporte em todos os fólios, em maior ou menor extensão. Tendo em vista esse cenário, é altamente recomendável, senão mandatório, fazer a leitura dessa documentação a partir de sua representação digital, criada, neste caso, a partir das fotografias das atas originais. Como os quatro primeiros livros de vereações passaram por um tratamento de conservação e restauro realizado em 2005 pela Fundação Vitae, conforme explicado em Monte & Fachin (a sair), atualmente cada fólio das atas encontra-se envolvido por um invólucro de poliéster, para proteger a documentação da umidade e garantir a conservação do suporte original, entre outras

---

<sup>2</sup> O *website* do Centro de Memória da Câmara, lançado recentemente, retorna os volumes impressos digitalizados em formato .pdf quando se pesquisam as atas antigas de Santo André e de São Paulo (cf. <http://www.camara.sp.gov.br/memoria>, Seção: Biblioteca).

funções. O trabalho de reprodução fotográfica, realizado no âmbito do Projeto *As atas da vila de São Paulo* (MONTE, 2015), teve como primeiro desafio, além da delicada manipulação dos originais, os obstáculos advindos do reflexo provocado pelo invólucro, como veremos a seguir. Pelo seu estado atual de fragilidade e considerando a tradicional etapa inicial do trabalho filológico, que é a leitura das fontes, a necessidade de se tratar digitalmente este material surge, em primeiro lugar, pelo imperativo de preservar os documentos originais: ao armazenar as informações sobre sua materialidade em formato eletrônico, pode-se minimizar a necessidade de seu manuseio. A digitalização como fator de preservação documental tem sido uma abordagem intensamente discutida nos anos recentes – veja-se, entre outros, Sayão et al. (2009) e Nardino et al. (2005). Alinhando-nos a esta abordagem, ressaltamos ainda que, entre as informações sobre a materialidade do documento a preservar no formato digital, importa incluir informações de natureza codicológica, diplomáticas e paleográficas.

Um segundo fator que leva a uma abordagem digital dessa documentação é a relevância de se proporcionar sua leitura a um público amplo constituído tanto por especialistas (linguistas, historiadores, sociólogos etc.) quanto pela comunidade em geral, interessada na representação de São Paulo em seus primeiros livros de vereação. Unindo a motivação do tratamento digital dos documentos (preservação) ao tratamento digital do texto (ampla difusão), surge o caminho de emprendermos o trabalho de edição filológica digital em sentido amplo a que denominamos *edição filológica virtual*: a produção de uma réplica informatizada e ressignificada do conjunto da documentação, detalhada a seguir.

## 2. Caminhos para uma edição filológica virtual

### 2.1 Conceito de edição filológica virtual

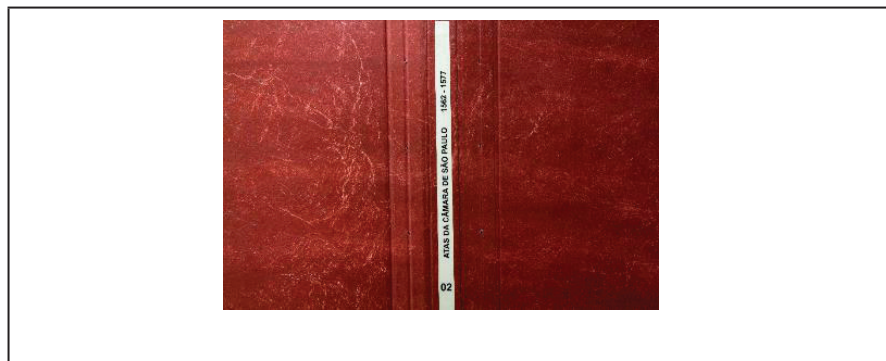
Usamos o termo *edição filológica virtual* para nos referir ao objeto criado a partir de um trabalho que inclui a produção da réplica imagética digital do documento físico, a edição filológica digital, e as múltiplas possibilidades de representação final ou publicação digital. Todos esses processos são construídos com ferramentas e tecnologias computacionais, e são, portanto, ‘digitais’; o conjunto dos processos forma o objeto que chamamos de virtual – pois que simula, representa, re-cria artificialmente os documentos originais. Produzir uma edição filológica virtual é empreender uma efetiva *tradução material* dos documentos, nos termos de Paixão de Sousa (2012): é digitalizá-los no sentido lato do termo, codificando-se as informações relevantes contidas no documento físico sob forma de símbolos computáveis. Almeja-se, neste processo, compor um novo objeto – o conjunto da documentação em formato digital – que ao mesmo tempo remete fielmente aos originais em papel, e proporciona possibilidades de pesquisa não imediatamente presentes nos originais, como a indexação para buscas remissivas, a anotação linguística, ou a inclusão de informações historiográficas.

Este novo objeto tem portanto três componentes básicos: a representação digital do texto original sob forma de imagem; a ‘textualização’ ou passagem da informação visual para informação efetivamente codificada como texto (incluindo-se a ‘transcrição’ do texto e a anotação de informações adicionais sobre o texto); e a apresentação do texto como objeto novamente legível como informação visual (de fato, num processo de *desmaterialização* e *rematerialização* do texto, de informação visual para informação computacional, e de volta para informação visual). Esses três componentes operam na construção da edição virtual das ACSP.

## 2.2 Etapas da edição

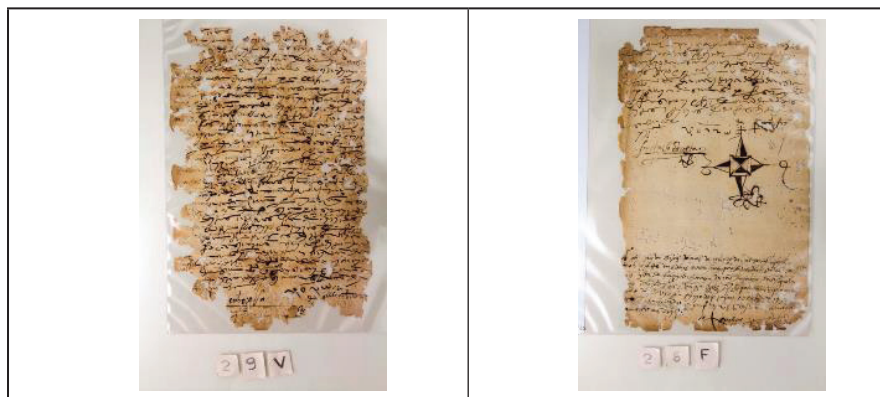
A construção da edição virtual tem início já com o processo de reprodução digital, ou ‘*digitalização*’: a captura e o processamento de imagens digitais a partir de objetos não-digitais. Este processo implica uma simulação da materialidade do objeto original, em particular quando o objetivo é realizar a leitura filológica a partir do documento digitalizado, e se busca construir imagens digitais que reflitam o mais fielmente possível as características físicas do original. De fato, já na captura de imagens digitais o objetivo da fidelidade ao original resulta de algum grau de manipulação das condições técnicas de captura. Na primeira etapa do Projeto “*As atas da vila de São Paulo*” (MONTE, 2015), foram realizados dois testes de fotografia para simular “*as condições ideais de iluminação, tempo de exposição, margens e, principalmente, rebatimento de luz para tentar minimizar o efeito reflexivo do invólucro de poliéster do documento*” (REGISTRO, 2016). No caso das ACSP, os documentos foram capturados em 300 Dpi-Raw (de ‘*cru*’, em inglês), e depois convertidos para JPEG (*Joint Photographic Experts Group*), pois o grande tamanho do formato Raw inviabilizaria o trabalho com os mesmos. Os *Quadros 1* e *2* mostram imagens em JPEG digitalizadas a partir do Volume 2 das ACSP:

QUADRO 1: Imagem em JPEG da Capa e Contra-capá do Vol. 2 das Atas de São Paulo





QUADRO 2: Imagem em JPEG dos fólios 26F e 29V do Vol. 2 das Atas de São Paulo

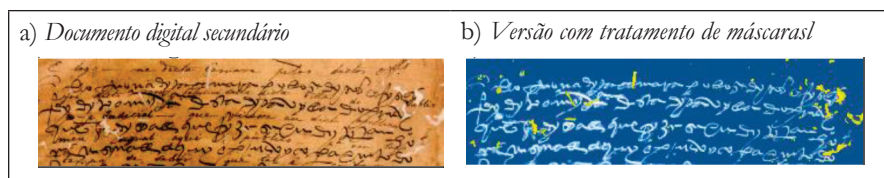


Considerando as múltiplas etapas de manipulação envolvidas na captura e pós-processamento, usamos aqui com o conceito de ‘*documento digital*’, proposto por Namiuti-Temponi et al. (2013), para descrever a reprodução digital obtida com métodos científicos de captura de imagem e codificação da informação. Chamamos a versão do documento resultante inicial da captura de imagens de ‘*documento digital primário*’ (neste caso, o arquivo Raw) e o documento pós-processado de *documento digital secundário* (neste caso, o JPEG). O documento digital secundário pode sofrer manipulações computacionais adicionais para melhorar as condições de leitura, trazendo benefícios notáveis à etapa de decifração do texto nas edições filológicas. Algumas técnicas neste sentido são propostas, por exemplo, por Havens (2014), usando o software proprietário Photoshop, e por Santos & Brito (2014), com um sistema desenvolvido pelos autores com base no conceito de ‘*documento digital*’ já mencionado. No caso das ACSP, realizamos testes de manipulação menos sofisticados<sup>3</sup>, mas já com um grau razoável de facilitação da leitura. É o caso, principalmente, dos fólios que apresentam duas camadas de escrita: a original quinhentista e a de um punho moderno, provavelmente de finais do século XIX,

<sup>3</sup> Com o software aberto GIMP 2.8.10, [www.gimp.org](http://www.gimp.org), 1997-2014.

início do XX, que ao decifrar o material, foi anotando no próprio a leitura feita. O *Quadro 3* abaixo mostra este fenômeno para o fôlio 32v do Volume 1, na Figura (a); a Figura (b) mostra uma versão de trabalho do fôlio, após um teste de manipulação para salientar a camada da escrita quinhentista e esconder a posterior:

QUADRO 3: Experimento de versões digitais para leitura



Vemos aqui revelar-se uma das particularidades importantes do *documento digital* em contraste com o documento em papel: o documento digital é apenas um código, e pode ser ‘clonado’ e manipulado sem prejuízo da versão original (isto é: do *documento digital primário*), mantida como referência ao longo do trabalho. Não estamos lidando com ‘o documento’, mas com múltiplas versões de uma representação artificial, descorporificada, de um documento.

Em que pesem as facilidades de leitura trazidas pela produção de documentos digitais como os ilustrados, no trabalho filológico e linguístico precisamos lidar com textos, não com imagens de textos. Entra, aqui, a segunda etapa do trabalho de edição: a transformação dos documentos digitais no formato imagem em documentos digitais no formato texto (no sentido computacional, um pequeno banco de dados que contém informações sobre codificação de caracteres, cf. Paixão De Sousa 2014). Para documentos digitais obtidos de originais impressos, essa transformação pode se dar por processamento automático; para manuscritos, a técnica segue sendo a decifração e transcrição humana – e no caso de materiais históricos como as ACSP, decifração humana por especialistas. Para além da decifração e transcrição, a edição filológica

inclui a anotação das informações relevantes sobre as decisões críticas do editor, uma etapa também passível de codificação eletrônica. É ao conjunto formado por esses dois processos (transcrição e anotação) somado ao seu produto em meio eletrônico que se denomina, em geral, ‘*edição filológica digital*’ – noção difusa, mas que tem tomado contornos conceituais mais definidos principalmente a partir de Crane, 2008; Gradmann & Meister, 2008; Lucía Megías 2008, 2009, 2010, a que referimos para um aprofundamento do debate conceitual.

Aqui cabe destacar as convergências e divergências entre o processo de edição no meio digital e as etapas do trabalho filológico tradicional. Podemos fazê-lo, seguindo Paixão de Sousa (2013b), partindo da clássica proposta de Segismundo Spina para as etapas básicas do trabalho de edição crítica - a *Fixação* do texto, ou seu preparo, apuração e estabelecimento; e a *Apresentação* do texto, ou “*organização técnica do texto e dos seus elementos elucidativos (...) tendo em vista sua publicação*” (SPINA, 1977: 86). Nesse sentido, a primeira diferença do trabalho de edição em meio digital em relação ao trabalho no meio não-digital é a radical separação entre essas duas etapas, formando-se dois objetos digitais distintos: um representando o texto ‘*fixado*’ e outro, o texto ‘*apresentado*’. Como aponta também Paixão de Sousa 2013b, na publicação em meio não-digital, a limitação do papel obriga o editor a tomar decisões que apagam alguns dos passos tomados durante a etapa de estabelecimento do texto, para criar uma apresentação final mais limpa e inteligível. Assim, “*não há **um objeto** que contenha todas as etapas do trabalho que este editor realizou sobre o texto*”. O trabalho com o documento eletrônico tem aqui uma segunda diferença técnica essencial: o editor pode deixar registrados todos os passos do estabelecimento do texto, anotando-os em único arquivo eletrônico, decidindo (ou deixando a cargo do leitor decidir) quais devem ou não estar explícitos na apresentação da edição. Assim, para Paixão de Sousa (2013b:16),

com a tecnologia digital, o elo entre o trabalho com o texto e sua apresentação final modifica-se. Já não há razão alguma para os “tipos de edição” serem tomados como objetos isolados entre si: podem, muito melhor, ser compreendidos como **camadas editoriais** possíveis sobre um mesmo texto – um mesmo texto lógico, aberto a múltiplas possibilidades de representação final. Ou seja, no contexto de uma edição digital, o trabalho de edição e o produto da edição desvinculam-se, uma vez que um mesmo texto pode conter anotações referentes a diversas etapas do trabalho de edição, reproduzíveis em diferentes camadas como produto final.

Esse conceito de edição em ‘*camadas*’ funda a metodologia de edição digital das ACSP que apresentamos a seguir. Partimos de um esquema simplificado de anotação (em formato texto simples, .txt), que passar a um formato mais complexo e flexível, permitindo a exploração das camadas de edição, como veremos.<sup>4</sup> Resumimos aqui este método com exemplos da edição semidiplomática de Monte & Fachin (a sair) para o fólio 37v das ACSP: o Quadro 4 mostra essa edição em (a), e sua anotação em (b).

---

<sup>4</sup> Essa etapa de anotação simplificada dispensa um conhecimento computacional muito extenso pelos editores. A marcação é também inteiramente independente de qualquer software ou ferramenta computacional: pode ser feita em um arquivo de texto simples (.txt), em qualquer editor de textos comum.

## QUADRO 4: Anotação simples de edição

- (a) Transcrição semidiplomática justalinear (Monte; Fachin, a sair)

Aos vinte eseis dias do mesde marco  
 de mill e quinhentos ese[t]enta e sinquo  
 anos nesta vjllade sampaullo nas pouasa  
 [d]as donde pouasa o vreador antonio cu[bas]  
 estando hi ã camara cõ seu parsei[ro]  
 antonio biquido e antonio preto juiz ordinario  
 fazendo efalando cousas nesesarjas para are

- (a) Anotação preparatória simples em .txt

```
_Aos_vinte_eseis_dias_do_mesde_marco_|
_de_mill_e_quinhentos_[eseenta][ese(t)enta]_e_sinquo_|
_anos_nesta_vjllade_sampaullo_nas_[pouasa$]
$as][pouasa$(d)as]_donde_pouasa_o_vreador_{anto}{ant(oni)o}_{cu}[cu(bas)]_|
_estando_hi_ã_camara_cõ_seu_[parsei][parsei(ro)]_|
_{anto}{ant(oni)o}_biquido_e_{anto}{ant(oni)o}_preto_juiz_ordinario_|
_fazendo_efalando_cousas_nesarjas_{pra}(a)ra_are$|
$publica_
```

Os símbolos usados nessa anotação (da série mais comum nos computadores pessoais) são marcas sobre as quais depois serão efetuadas substituições automáticas gerando uma anotação mais complexa. O traço baixo \_ indica espaço entre palavras<sup>5</sup>; |, quebra de linha; \$, quebra de palavra; [...][...], sequências de palavra original e editada, nos casos de conjecturas; {...}{...}, sequências de palavra original e editada com desenvolvimento de abreviaturas; (...), o caracter acrescentado, nos dois casos. Por exemplo, em 4(a) ‘[eseenta][ese(t)enta]’ indica que a lição original é ‘eseenta’, e a intervenção do editor por conjectura é ‘ese(t)enta’ (i.e.: adição do ‘t’); ‘{anto}{ant(oni)o}’ indica que a edição diplomática é

<sup>5</sup> Neste ponto caber um esclarecimento sobre a marcação da separação vocabular com um caractere visível (e não, simplesmente, um ‘espaço’, que, no texto digital, equivale a um caractere invisível). De um lado, a separação vocabular pode ser uma informação relevante para futuros estudos linguísticos; de outro, a separação por espaços é sujeita a erros humanos e problemas de decodificação na passagem entre programas de processamento de texto; esta anotação garante menor incidência de enganos por parte do digitador e corrupção lógica. Na passagem do arquivo simples para a anotação final em XML, cada ‘palavra’ (no sentido de ‘palavra gráfica’) será codificada pelo elemento <w>...</w>, como veremos.

‘*anto*’, e o desenvolvimento de abreviaturas é ‘*ant(oni)o*’ (i.e., adição de ‘*oni*’). Esta anotação, apesar de conveniente, não pode ser a anotação final dos documentos: para que a edição seja legível em larga escala, e compartilhada no âmbito acadêmico (e mesmo para além dele) via internet, faz-se necessária uma linguagem comum adotada por diferentes projetos e legível em diferentes navegadores.

Nesse sentido, a principal opção é a linguagem XML (*eXtended Markup Language*<sup>6</sup>, W3C 2016), por sua combinação equilibrada entre flexibilidade e controle, e por ser a base da maioria dos corpora eletrônicos e edições filológicas digitais atuais<sup>6</sup>. No âmbito da língua portuguesa, o primeiro projeto a usar essa linguagem para edições com finalidade acadêmica foi o *Corpus Tycho Brahe*<sup>7</sup>, com o sistema proposto em Paixão de Sousa & Trippel (2004, 2006), Paixão de Sousa et al. (2010), detalhado em Paixão de Sousa (2014), e atualmente usado por diferentes projetos brasileiros e portugueses. Para a edição das ACSP, a anotação simples de base (cf. *Quadro 4*) gera um código XML nos moldes desse sistema, como no *Quadro 5*:

QUADRO 5: Anotação XML da edição semidiplomática (20 primeiras palavras)

```
<W>Aos</W>
<W>vinte</W>
<W>eseis</W>
<W>dias</W>
<W>do</W>
<W>mesde</W>
<W>marco</W><br/>
<W>de</W>
<W>mill</W>
<W>e</W>
<W>quinhentos</W>
<W><or>esenta</or><ed t="cjt">ese<ed t="add">t</ed>enta</ed></W>
<W>e</W><W>sinquo</W><br/>
<W>anos</W>
<W>nesta</W>
<W>vjllade</W>
<W>sampaullo</W>
<W>nas</W>
<W><or>pousa<br/>as</or><ed t="cjt">pousa<br/><ed t="add">d</ed>as</ed></W>
<W>dondes</W>
<W>pousa</W>
<W>o</W>
<W>vreador</W>
<W><or>anto</or><ed t="exp">ant<ed t="add">oni</ed>o</ed></W>
<W><or>cu</or><ed t="cjt">cu<ed t="add">bas</ed></ed></W><br/>
```

<sup>6</sup> Em especial, graças à *Text Encoding Initiative* (TEI), consórcio dedicado à standardização e manutenção da codificação de textos eletrônicos com fins acadêmicos, cf. <http://www.tei-c.org/index.xml>

<sup>7</sup> Cf. <http://www.tycho.iel.unicamp.br/corpus/>

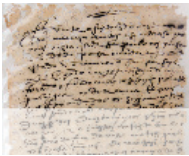
Com esta técnica, um mesmo arquivo anotado contém todas as informações do texto original consideradas relevantes pelo editor e todas as interferências realizadas pelo editor. As camadas editoriais assim registradas são passíveis de diferentes representações finais para leitura, automaticamente geradas com a seleção de uma ou outra forma anotada no arquivo de base, na etapa que segue a digitalização, transcrição e anotação de edição – isto é, a publicação.

A ‘publicação’, de um ponto de vista computacional, equivale ao momento da recomposição do documento em forma de informação visual: opera aqui a simulação de um texto legível a partir dos códigos computacionais incutidos nas etapas anteriores. Pensamos aqui portanto fundamentalmente na publicação *digital* do resultado dos processos anteriores. Guarde-se em mente, entretanto, a equivalência com a etapa da ‘*Apresentação do texto*’ em Spina (1977), malgrado algumas diferenças importantes, já apontadas acima: o sistema de anotação ora apresentado permite a preservação das diversas etapas de interferência sem prejuízo da possibilidade de inúmeras formas de apresentação. No arquivo de base, a transcrição mais próxima ao original está reunida à transcrição com mais intervenções do editor, e as formas de apresentação possível irão selecionar as camadas de informação a serem extraídas, conforme o destino do texto. Vale dizer, portanto, que o arquivo de anotação contém as informações que se *transformarão* na “*edição diplomática*”, na “*edição semidiplomática*” ou na “*edição modernizada*” (melhor dito, nas *apresentações* “*diplomática*”, “*semidiplomática*” ou “*modernizada*”) ou em catálogos de variações, glossários, e outras versões possíveis.

Assim, e ainda a partir o mesmo trecho do fôlio 37v do vol. 2 das ACSP, mostramos algumas das apresentações que poderiam ser geradas a partir da anotação XML básica já mostrada no *Quadro 5*. A apresentação no *Quadro 6* abaixo pode ser gerada a partir daquela anotação por meio de transformações XSLT (*eXtended Language for Stylesheet Transformation*, W3C 2015), gerando-se um arquivo HTML (*Hipertext Markup Language*,

W3C 2016b) – ou seja, um hipertexto, com a aparência ilustrada na coluna da esquerda:

QUADRO 6: Apresentação da edição semidiplomática a partir do arquivo XML

<p><i>Aos vinte e seis dias do mesde marco de mill e quinhentos e[se]nta e sinquo anos nesta vjllade sampaullo nas poua [d]as donde pouosa o vreador antonio cu[bas] estando hi e camara cõ seu parsei[ro] antonio biquido e antonio preto juiz ordinario fazendo efalando cousas nesesarjas para are</i></p>	
---	---

Um ponto interessante se faz ver neste passo: a ilustração no *Quadro 6* parece idêntica à transcrição semidiplomática de Monte & Fachin (a sair), já mostrada no *Quadro 4*. Essa semelhança, porém é fruto de artifício, e esconde um extenso processamento computacional de informação. Agora, temos um texto que **imita** a aparência de uma transcrição semidiplomática tradicional, mas que é produzido por sucessivas transformações automáticas a partir de um código-base. A primeira vantagem deste processo é justamente a de que por trás da representação assim criada há um texto efetivamente composto como um banco de dados compartilhável e processável; a segunda vantagem são as formas diferenciadas de apresentação que ele pode gerar, como veremos a seguir.

### 2.3 Para além da edição filológica tradicional

Como ressaltamos logo acima, a técnica de edição em camadas pode gerar representações muito próximas, visualmente, dos produtos da edição filológica tradicional. Este é em si um resultado interessante, pois nos permite trabalhar em um meio flexível e controlável, sem prejuízo de uma representação de texto com a qual a comunidade acadêmica está familiarizada. Podemos porém ir além dos formatos finais tradicionais,



construindo anotações e apresentações de textos particulares ao modo de difusão digital.

Aqui mostramos algumas dessas possibilidades, começando por técnicas já testadas em diferentes projetos. Nesse sentido, ressalte-se que depois de realizada a transcrição semidiplomática do texto nos moldes acima, o mesmo arquivo pode ser editado em camadas adicionais, por exemplo, modernizando-se as grafias. Esta técnica tem sido seguida longamente no *Corpus Tycho Brahe* (ambiente no qual as versões modernizadas dos textos são base para uso de outras ferramentas computacionais), e segue os mesmos moldes da anotação que codifica a intervenção na edição semidiplomática. No *Quadro 7*, (a) mostra a atualização de ‘*efalando*’ para ‘*e falando*’<sup>8</sup>; e (b) mostra a apresentação da edição modernizada, no mesmo espírito do que foi mostrado acima para a edição semidiplomática.

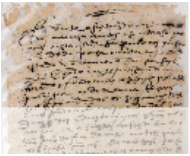
#### QUADRO 7: Camadas editoriais adicionais: Anotação de modernização de grafia

(a) Anotação XML:

```
<w>efalando</w> =
<w><or>efalando</or><ed t="dis">e falando</ed></w>
```

(b) Apresentação da edição modernizada simples

Aos vinte e seis dias do mês de março de mil e quinhentos e setenta e cinco anos nesta vila de São Paulo nas pousadas donde pousa o vereador Antonio Cubas estando aí em câmara com seu parceiro Antonio Bicudo e Antonio Preto juiz ordinário fazendo e falando coisas necessárias para a república...



<sup>8</sup> O elemento ‘<w>...</w>’ remete à palavra ‘*gráfica*’, ou seja, ao espaçamento do documento original (mais que às palavras morfológicas interpretadas pelo editor). O elemento ‘<or>...</or>’ remete à transcrição mais conservadora, que preserva a fronteira de palavras original, e ‘<ed>...</ed>’ à versão modernizada.

Além da facilitação da leitura (lembrando-se o objetivo de ampliar o público leitor desta documentação), a versão modernizada pode fundar trabalhos que enriqueçam o texto com informações de ordens diversas, como informações biográficas sobre as figura históricas e contextualização histórico-geográfica dos topônimos mencionados<sup>9</sup>. O *Quadro 8* mostra uma anotação com informações historiográficas em (a), e uma apresentação da anotação em (b):

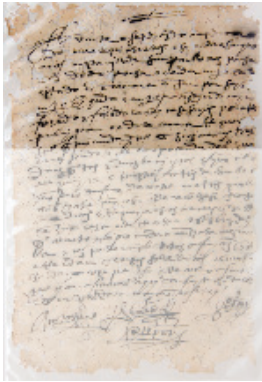
QUADRO 8: Enriquecimento com informações historiográficas (de LEME, 1905)

(a) Anotação XML	<pre>&lt;referente id="0002" nome="Antonio Bicudo" entrada="Antonio Bicudo Carneiro foi da governança da terra, ouvidor da comarca e capitania pelos anos de 1585; foi quem mandou levantar pelourinho na vila de S. Paulo no dito ano de 1585. Foi casado com Izabel Rodrigues, natural de S. Paulo, segundo se vê do testamento de seu f.º Antonio Bicudo em 1659 em que declara sua filiação. Teve 6 f.ºs, como se vê do requerimento feito aos oficiais da câmara de S. Paulo pedindo em 1598 chãos para fazer casas" fonte="G.P., VI, 296"&gt;&lt;w&gt;&lt;or&gt;&lt;ed t="exp"&gt;ant&lt;ed t="add"&gt;ont&lt;/ed&gt;&lt;/w&gt;&lt;/w&gt;&lt;bicudo&lt;/w&gt;&lt;/referente&gt;</pre>
(b) Apresentação	<p>Aos vinte e seis dias do mês de março de mil e quinhentos e setenta e cinco anos nesta vila de São Paulo nas pousadas donde pousa o vereador Antonio Cubas estando aí em câmara com seu parceiro <a href="#">Antonio Bicudo</a>, e Antonio Preto juiz ordinário fazendo e falando coisas necessárias para a re-</p> <div style="border: 1px solid black; padding: 5px; width: fit-content; margin-left: auto; margin-right: auto;"> <p>Antonio Bicudo Carneiro foi da governança da terra, que do ra comarca e capitania pelos anos de 1585; foi quem mandou levantar pelourinho na vila de S. Paulo no dito ano de 1585. Foi casado com Izabel Rodrigues, natural de S. Paulo, segundo se vê do testamento de seu f.º Antonio Bicudo em 1659 em que declara sua filiação. Teve 6 f.ºs, como se vê do requerimento feito aos oficiais da câmara de S. Paulo pedindo em 1598 chãos para fazer casas" [G.P.]</p> <p>[Ver outras ocorrências de Antonio Bicudo nas atas]</p> </div>

É possível, também, combinar-se a anotação de mais de um texto nos mesmos moldes, gerando-se apresentações paralelas. Seria possível, por exemplo, combinar a transcrição semidiplomática das ACSP e a transcrição da edição impressa de 1967 (aquela que, como apontamos, tornou-se sinônimo das *Atas*), gerando uma apresentação com as duas versões alinhadas, como no *Quadro 9*:

<sup>9</sup> É claro que se trata necessariamente de um trabalho multidisciplinar; no caso específico das ACSP, os historiadores seriam os especialistas mais indicados para empreendê-lo. No exemplo que segue, como experimento, usamos a entrada sobre Antonio Bicudo na *Genealogia Paulistana* (LEME, 1905).

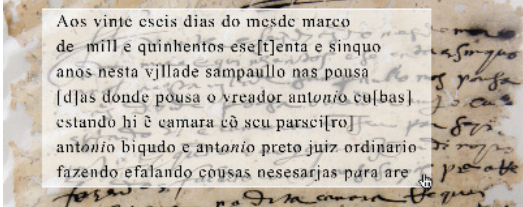
QUADRO 9: Apresentação de edição diplomática justalinear paralela à edição de 1967

<p>Aos vinte e seis dias do mesde marco <i>Aos vinte e seis dias do mes de marco</i></p> <p>de mill e quinhentos ese[t]enta e sinquo <i>de mill e qui- nhentos e setenta e sinquo</i></p> <p>anos nesta vjllade sampaullo nas pouasa <i>anos nesta villa de sam paullo   nas pousadas</i></p> <p>[d]jas donde pouasa o vreador antonio cu[bas] <i>donde pouasa o vreador ant<sup>o</sup> cubas</i></p> <p>estando hi ã camara cõ seu parsei[ro] <i>estando  hi ã camara cõ seu parseiro</i></p> <p>antonio biquido e antonio preto juiz ordinario <i>ant<sup>o</sup> biquido e ant<sup>o</sup> preto  juiz ordinario</i></p> <p>fazendo efalando cousas nesesarjas para are <i>fazendo e falando cousas nesesarria pr<sup>o</sup>   a republicua</i></p>	
--	--

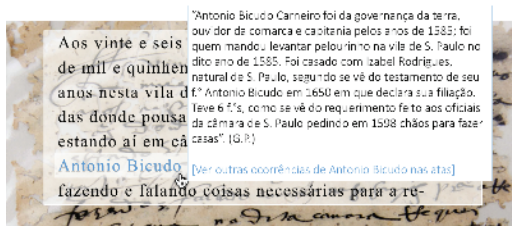
As diferentes apresentações até este ponto mostram o fac-símile do documento em paralelo ao texto. Pode-se, ainda, sobrepor-se o texto editado (em qualquer versão) ao fac-símile, com apresentações que se iniciam com a visualização apenas do fac-símile, ativando-se o texto decifrado ao comando do leitor, como no *Quadro 10*:

QUADRO 10: Apresentações sobrepostas

(a) fac-símile com ativação de sobreposição de edição



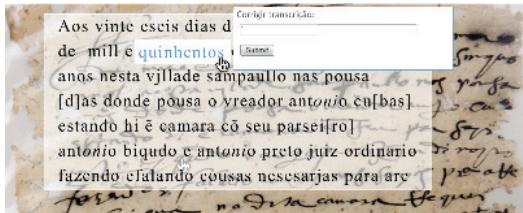
(b) fac-símile com ativação de sobreposição de edição modernizada e texto enriquecido



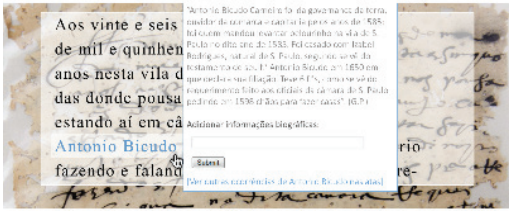
Estes exemplos nos remetem a um ponto interessante, e de particular relevância no caso das ACSP. Como já mencionamos, há duas camadas de escrita na documentação: a original, quinhentista, e a moderna, dos punhos que decifraram as atas. Veja-se então que, com a edição virtual, é possível replicar a intenção do leitor-editor que adicionou sua transcrição ao texto original (supomos, para facilitar a decifração), sem incorrer na mesma consequência pouco desejável (a alteração definitiva do documento). Podemos agora sobrepor nossa leitura ao documento sem prejuízo do mesmo: pois não estamos de fato incorporando nada a este objeto que, como insistimos, não tem corpo. Com efeito, tal é a flexibilidade permitida pela separação entre o objeto físico e o objeto publicado na edição virtual que nem seu resultado final é imutável. Podemos construir versões interativas da nossa edição, possibilitando a interferência do leitor na forma apresentada, sugerindo novas leituras, corrigindo eventuais enganos de decifração, ou enriquecendo a camada de informações historiográficas. É o que sugere o *Quadro 11* abaixo.

QUADRO 11: Apresentações interativas

(a) Apresentação com abertura para intervenção na transcrição



(a) Apresentação com abertura para enriquecimento da informação historiográfica



Estamos aqui de fato diante do documento ‘descorporificado’, fragmentado, pulverizado, rematerializado. E embora esses termos nos possam causar um arrepio involuntário frente à impressão de alteração e desfiguração de uma documentação tão rara, na realidade o contrário acontece: manipulamos o texto intensamente, sem nada alterar o objeto original que nos trouxe o texto. Este é, a nosso ver, o impacto mais claramente positivo da edição virtual.

**Conclusão**

Sugerimos, neste artigo, o termo **edição filológica virtual** para descrever o trabalho de edição iniciado com as ACSP. Esta escolha remete inicialmente às consequências práticas deste trabalho de edição em meio digital: fundamentalmente, queremos destacar que

este trabalho de manipulação do texto – em certos pontos extremo – não tem nenhum impacto material negativo sobre este conjunto documental raro e frágil. O trabalho produz, justamente, uma simulação dos documentos, uma edição *virtual* pois que paralela à existência dos objetos de papel – que podem, assim, permanecer sob a proteção que merecem frente à manipulação dos leitores. Esta é uma característica de todo trabalho de digitalização de fontes (de um lado) e de edição digital (de outro); e pensamos que ao unir essas duas etapas num mesmo objeto digital, estamos apenas potencializando os efeitos positivos da preservação. Trata-se entretanto de uma preservação democratizada: garante-se com ela a um tempo a integridade física do documento e sua ampla difusão para o público leitor, como é nosso objetivo. Estas são, portanto, as principais consequências práticas da edição virtual. De outro lado, no campo teórico, como resultado do experimento inicial com essa documentação que encerra tantos desafios filológicos, vemos consolidarem-se os conceitos sugeridos na reflexão de Paixão de Sousa (2013a) no sentido de tomar como aspecto central da natureza do texto digital sua propriedade **descorporificada**. É por conta da dissipação do elo entre o objeto produzido e o objeto finalmente lido que as vantagens ‘práticas’ da edição virtual, acima apontadas, podem de fato se dar. É apenas porque o texto digital é *algo que não se pode segurar com as mãos*<sup>10</sup> que ele se torna a melhor forma de se conservar, estudar e difundir a preciosa documentação aqui inicialmente trabalhada – e sobre a qual, certamente, muito ainda é necessário (e desejável) trabalhar. Esperamos, com o experimento e a reflexão inicial aqui apresentados, ter apontado caminhos interessantes para esses futuros trabalhos.

---

<sup>10</sup> Como diz Paixão de Sousa 2013a, em paráfrase à colocação de Carter (1968) sobre a natureza material do texto impresso: “*Type is something you can pick up and hold in your hands?*”.

## Referências

CARTER, H. **A view of early typography**. [The Lyell Lectures, 1968. Oxford: Clarendon Press, 1969]. Ed rev. Londres: Hyphen Press, 2002.

CHAUDIRON, S.; IHADJADENE, M.; MAREDJ, A. **La fragmentation et l'unité documentaire en question**, Actes du 16ème Congrès de la SFSIC, 11-13 juin 2008, Compiègne.

CLEMENT, T. **Knowledge Representation and Digital Scholarly Editions in Theory and Practice**. Journal of the Text Encoding Initiative, 1, June 2011.

CRANE, G. (et al.). **ePhilology: when the books talk to their readers**. In: S. Schreibman, R. Siemens (eds), *A Companion to Digital Literary Studies*. Oxford: Blackwell, 2008.

EGGERT, P. *Text as Algorithm and as Process*. In W. McCarty (Ed.), *Text and Genre in Reconstruction: Effects of Digitalization on Ideas, Behaviours, Products and Institutions*. Cambridge: Open Book, 2010 (183-202).

GRADMANN, S.; MEISTER, J. C. **Digital document and interpretation: rethinking "text" and scholarship in electronic settings**. *Poiesis Prax*, 5:139–153., 2008.

HAVENS, H. **Adobe Photoshop and Eighteenth Century Manuscripts: A New Approach to Digital Paleography**. *Digital Humanities Quarterly*, V. 8 N. 4, 2014.

HOUAISS, A. **Grande Dicionário Houaiss online**. 2012.

LEME, L. G. da S. **Genealogia Paulistana**. São Paulo: Duprat, 1905. Vol VI, p. 297.

LUCÍA MEGÍAS, J. M. **De las bibliotecas digitales a las plataformas de conocimiento (notas sobre el futuro del texto en la era digital).**

In: Estudos de edición crítica e lírica galego-portuguesa. Verba. Anuario galego de Filoloxía. Anexo 67. Universidade de Santiago de Compostela, Santiago de Compostela: 2010. p. 369-401.

\_\_\_\_\_. **La edición crítica hipertextual.** In: Lecturas y textos en el siglo XXI: los nuevos caminos de la crítica textual. AXAC, Lugo, 2009.

\_\_\_\_\_. **Las relaciones entre la bibliografía textual y la informática humanística: el incunable del hipertexto.** Tipofilología. Rivista Internazionale di Studi Filologici e Linguistici sui testi a stampa (1), 2008, pp. 119-138.

MONTE, V. M. do; FACHIN, P. R. M. **Notícia da edição filológica das atas de vereações quinhentistas de São paulo e de Santo André da Borda do Campo.** A sair.

MONTE, V. M. do. **As atas da vila de São Paulo.** Projeto de Pesquisa, Programa Unificado de Bolsas (PUB), Reitoria da Universidade de São Paulo, 2015.

\_\_\_\_\_. **As atas da vila de Santo André da Borda do Campo e de São Paulo: reprodução digital e edição.** Projeto de Pesquisa, Programa Unificado de Bolsas (PUB), Reitoria da Universidade de São Paulo. 2016b.

\_\_\_\_\_. Simpósio Arquivo Histórico de São Paulo: Memórias, Práticas e Perspectivas – Mesa **Caminhos de Pesquisa no Arquivo Histórico de São Paulo: Uma Visão Multidisciplinar.** São Paulo, 06/10/2016. 2016a.

NAMIUTI-TEMPONI, C.; SANTOS, J. V.; COSTA, A.; FARIAS, I. S. **Computação e Linguística: importante diálogo para pesquisas e preservação da memória nos novos meios das antigas fontes.** RBBA, v. 2, n. 1, p. 9-34, 2013.



NARDINO, A. T. D.; CAREGNATO, S. E. **O futuro dos livros do passado: a biblioteca digital contribuindo na preservação e acesso às obras raras.** Em *Questão*, Porto Alegre, v. 11, n. 2, p. 381-407, jul./dez. 2005.

PAIXÃO DE SOUSA, M. C. **A Filologia Digital em Língua Portuguesa: Alguns caminhos.** In: BANZA, A. P.; GONÇALVES, M. F. (coord.), *Património textual e humanidades digitais: da antiga à nova Filologia.* Évora: Centro Interdisciplinar de História, Culturas e Sociedades da Universidade de Évora (CIDEHUS)/ Fundação para a Ciência e a Tecnologia (FCT), 2013(b).

\_\_\_\_\_. **Conceito material de texto digital: Um ensaio.** *Texto Digital*, v. 2009, p. 6, 2009.

\_\_\_\_\_. **O Corpus Tycho Brahe: contribuições para as humanidades digitais no Brasil.** *Filologia e Linguística Portuguesa*, v. 16, p. 53-93, 2014.

\_\_\_\_\_. **O livro impresso raro e o ambiente de leitura digital: um debate com base no espólio dos Craesbeeck.** *Seminário Mindlin 2012: Um mundo sobre papel: livros, gravuras e mapas na era dos descobrimentos.* São Paulo, 2012.

\_\_\_\_\_. **Texto digital: Uma perspectiva material.** *Revista ANPOLL (Associação Nacional de Pós-Graduação e Pesquisa em Letras e Linguística).* Volume 1, Número 35, 2013 (a).

PAIXÃO DE SOUSA, M. C.; KEPLER, F. N.; FARIA, P. P. F. ***E-Dictor: Novas perspectivas na codificação e edição de corpora de textos históricos.*** In: T. Shepherd; T.B. Sardinha; M.V. Pinto. (Org). *Caminhos da linguística de corpus.* Campinas: Mercado de Letras, 2010.

PAIXÃO DE SOUSA, M. C.; TRIPPEL, T. **Building a historical corpus for Classical Portuguese: some technological aspects.** V International Conference on Language Resources and Evaluation (LREC) 2006: Proceedings, 2006.

REGISTRO, T. **Relatório de Reprodução Fotográfica das Atas de São Paulo**. São Paulo, 2016.

SANTOS, J. V.; BRITO, G. S. **Fotografia técnica de documentos para formação de corpora digitais eletrônicos: o método desenvolvido no Lapelinc**. Letras & Letras, v. 30, n. 2, 2014.

SAYÃO, L.; TOUTAIN, L. B.; ROSA, F. G.; MARCONDES, C. H. (Orgs.). **Implantação e gestão de repositórios institucionais : políticas, memória, livre acesso e preservação**. Salvador : EDUFBA, 2009.

SPINA, S. **Introdução à Edótica**. São Paulo: Cultrix/EDUSP, 1977.

TRIPPEL, T.; PAIXÃO DE SOUSA, M. C. **Single source processing of Historic corpora for diverse uses**. Association for Literary and Linguistic Computing (ALLC) Annual Conference, Göttingen, 2004.

UNSWORTH, J. **Forms of Attention: Digital Humanities Beyond Representation**. The Face of Text: Computer-Assisted Text Analysis in the Humanities, III Conference of the Canadian Symposium on Text Analysis (CaSTA), McMaster University, November 19-21, 2004.

\_\_\_\_\_. **Knowledge Representation in Humanities Computing**. National Endowment for the Humanities, 2001.

W3C (a). *XML*. <http://www.w3.org/XML>, 2016.

W3C (b). *XSLT*. <https://www.w3.org/Style/XSL>, 2015.

W3C (c). *HTML*. 2013. <https://www.w3.org/html>, 2016.

Recebido em 30/09/2016 e aceito em 06/12/2016