

Abordagens digitais para estudos de Paleografia: desafios, atualidade, desdobramentos

Digital approaches to Paleography studies: challenges, actuality,
developments

DOI: <https://doi.org/10.24206/lh.v7i3.42771>

Antonio Ackel

Doutorando em Filologia e Língua Portuguesa pela Universidade de São Paulo e pela Universiteit Leiden com bolsa CAPES/PrInt (2021). Mestre pelo mesmo programa brasileiro com estágio na Rijksuniversiteit Groningen (2019) com bolsas Erasmus+ ICM, CNPq e CAPES (2018-2019). Pós-graduação em Docência no Ensino Superior pela Pontifícia Universidade Católica do Rio Grande do Sul (2018). Pesquisador visitante na Katholieke Universiteit Leuven (2016). Graduação em Linguística pela Universidade de São Paulo (2016). Campos de pesquisa: Filologia, Paleografia e História da Língua Portuguesa.

E-mail: antonio.ackel@usp.br

ORCID: <https://orcid.org/0000-0002-8283-4417>

RESUMO

Este texto busca resgatar fatos históricos que levaram a Paleografia Digital a se consubstanciar como uma disciplina, com teorias e métodos, dentro do campo das Humanidades Digitais. Apresenta também o estado da arte dessa disciplina, apontando testes e hipóteses que foram propostos ao longo dos anos. Como exemplo de uso de ferramenta tecnológica computacional na análise paleográfica, apresenta-se uma possível técnica para que textos possam ser lidos sob rasuras, por meio do programa *Sketchbook*. Para o estudo de caso, utilizou-se um datiloscrito com intervenções, à mão, de sua autora, Lygia Fagundes Telles, ao revisar sua obra *Seminário dos Ratos*, lançada em 1977. Após a apresentação de resultados parciais que vêm sendo obtidos com essa técnica, são feitas algumas considerações finais sobre os desafios que a paleografia ainda tem pela frente.

Palavras-chave: Humanidades Digitais. Paleografia Digital. Documentos históricos. Edição de imagens. Rasuras.

ABSTRACT

In this work we aim to rescue historical facts that led Digital Palaeography to become a discipline with theories and methods, within the field of Digital Humanities. We also present the state of art of it with tests and hypotheses that authors have been proposing over the years. As an example of a computational technological tool in paleographic analysis, we consider a possible technique that makes the text visible under erasures, using *Sketchbook* software. A Lygia Fagundes Telles' typewritten is used as test in this experience. It refers to her book *Seminário dos Ratos*, published in 1977. We finish our work thinking about challenges that digital paleography still faces in the future.

Keywords: Digital Humanities. Digital Palaeography. Historical documents. Image editing. Erasures.

Introdução

O objetivo deste artigo é aumentar o embasamento teórico já consolidado da Paleografia Digital (CIULA, 2005), oferecendo uma proposta com utilização de programas computacionais de edição de imagens. A exemplo de trabalhos como os de Palma *et al.* (2010) e Havens (2014), aqui, busca-se decifrar fragmentos textuais que estão sob rasuras.

O material utilizado como exemplo é um datiloscrito da obra *Seminário dos Ratos*, de Lygia Fagundes Telles (1977)¹, em que a autora faz diversas interferências, detalhado mais adiante.

Pode-se dizer que a utilização de recursos tecnológicos nos estudos da Paleografia não é de hoje, se considerarmos as técnicas, para a época, avançadas, de impressão que Jean Mabillon (1704) aplicou em sua obra fundamental, *Librorum de re diplomatica supplementum*, ou mesmo os recursos aplicados às imagens das pranchas de facsímiles e inscrições de Bond, Thompson e Warner (1884), nos dois primeiros volumes de *The Palaeographical Society*.

Na onda de novas tecnologias, recorda-se da proposta do desenvolvimento da escrita de Jean Mallon, nos anos de 1950, que foi apresentada em 1976, sob forma de cinema, com um filme dirigido por Jean Venard (FRAUZAURIH, 2016). Como exemplo de criação de banco de dados para análises codicológicas quantitativas, pode-se citar Gilissen (1977), que aplicou medidas estatísticas em seu estudo da escrita.

O trabalho de Albert Smith (1938) também deve ser considerado pela utilização de recurso fotográfico para recuperação de escrita ilegível em manuscritos. Seu método para ler textos apagados ou danificados empregava câmeras que produziam, e ainda produzem, imagem multispectral². Vejam-se os trabalhos de Chabries, Booras, Bearman, (2013) e de Edwards *et al.* (2019). O problema desse método, até hoje, é o alto custo para utilização.

O rol de autores e suas propostas para usos tecnológicos nos estudos paleográficos poderiam continuar em uma longa lista neste texto, no entanto, já fica claro o interesse que o homem sempre teve em utilizar-se da máquina para criar teorias e métodos que visassem ao entendimento correto de um texto escrito em qualquer superfície.

Diante dessa primeira observação, coloca-se a questão de como situar a epigrafia em ambiente digital, já que os estudos paleográficos dessa forma também são utilizados em suportes rígidos.

¹ O ano serve apenas como referência à primeira edição. O manuscrito não é datado, mas presume-se que sua elaboração tenha sido próxima à publicação da obra.

² A imagem multi-espectral foi desenvolvida para explorar as superfícies geográficas e foi adaptada para ler e registrar manuscritos danificados (cf. EDWARDS *et al.*, 2018).

O termo paleografia digital ou computacional foi atribuído pela primeira vez no trabalho de Ciulla (2005), que propôs novas perspectivas quanto ao seu significado, nomenclaturas e aspectos históricos. Sua teoria baseia-se no modelo semântico³ que se estrutura em dois eixos: imagem e significado (CIULLA, 2005, p. 100).

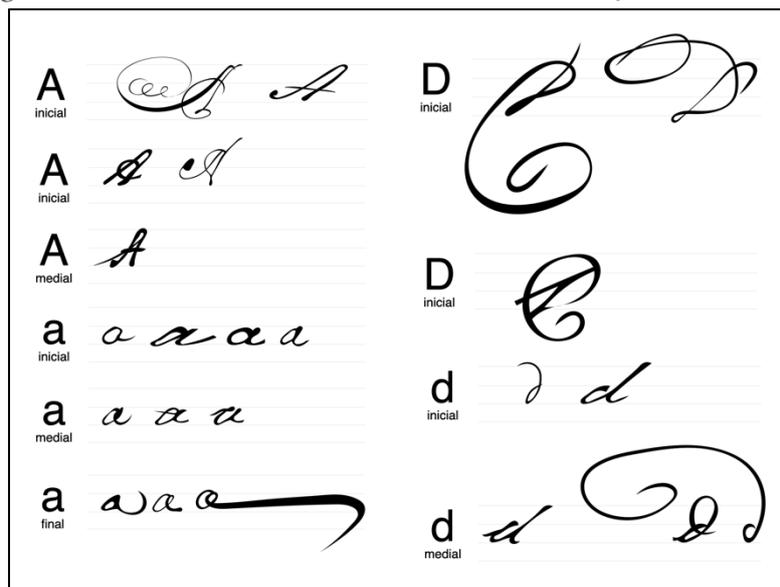
Dando continuidade aos estudos de Ciulla (2005), Hassner *et al.* (2013) avaliaram a união entre paleografia e computação e, por meio de associações entre imagem e significado do texto, criaram um grande banco de dados. Questões sobre bancos de dados devem ser consideradas, pois trata-se de um dos primeiros passos a serem tomados na análise paleográfica digital e, por essa razão, também serão abordados neste texto.

Das propostas sobre os estudos paleográficos em ambiente digital que têm sido propostos e avaliados ao longo dos últimos cinco anos, foi criado por nós um banco de dados de identificação de diferentes tipos de escrita intitulado, *Alfabetos do IEB/USP*, cujas formas das letras podem ser observadas em manuscritos, em língua portuguesa, que vão dos séculos XV ao XX. Ainda em fase de consolidação, alguns documentos, que foram usados para elaboração dos alfabetos, estão disponíveis em bibliotecas digitais e outros, apenas presencialmente. Trabalhos sobre essas formas gráficas podem ser consultados em Ackel (2019, 2021) e Ackel e Módolo (2020).

Cedo para que seja considerado um projeto, o estudo *Alfabetos do IEB/USP* é constituído por aproximadamente 8 mil símbolos gráficos, dentre letras, palavras e abreviaturas. No entanto, todos esses elementos estão disponíveis no manuscrito, ou seja, são claramente visíveis e podem ser lidos sem interferências na imagem. Abaixo, apresenta-se um modelo da catalogação das letras maiúsculas e minúsculas, *a* e *d*. As formas foram retiradas de um códice sem data, cuja autoria é atribuída ao Padre Antonio Vieira.⁴

³ Tradução nossa de “*semantic model*”.

⁴ O códice foi utilizado como exemplo por fazer parte da constituição do estudo de alfabetos já citado. Está disponível para consulta sob o código de referência AL-IEB/USP-138 e armazenado na caixa 06, 01, da coleção Alberto Lamego, do Instituto de Estudos Brasileiros da Universidade de São Paulo.

Figura 1 – Letras *a* e *d*, retiradas do AL-IEB/USP-138 da Coleção Alberto Lamego

Fonte: Elaboração própria.

É claro que a consulta a esses alfabetos colabora como fonte de pesquisas para aqueles que necessitam fazer comparações de símbolos gráficos semelhantes.

Mas, pensando-se em ferramentas virtuais que a paleografia digital oferece, procurou-se encontrar um método que melhorasse a visualização de fragmentos textuais danificados no suporte, fazendo uso de computadores residenciais, em comparação às máquinas de leitura multiespectrais.

Dessa forma, é apresentada aqui uma proposta de desenvolvimento técnico de apagamento de elementos que não fazem parte do texto escrito pelo autor, com a utilização do programa computacional *Sketchbook* (SKETCHBOOK, 2021). Esse programa tem sido base para outros trabalhos de edição de imagem e tem se mostrado satisfatório, como será apresentado em seção pertinente.

Aqui, como elemento não-textual-autoral, mostra-se uma forma de apagar traços de rasura sobre o texto, como dito, em um fac-símile correspondente ao datiloscrito IMS/LFT-08.014, da obra *O Seminário dos Ratos*, de Lygia Fagundes Telles (IMS, 2018).

A razão da escolha desse material se dá em função de algumas características. Em primeiro lugar, considera-se o fato de os manuscritos de Lygia serem de fácil acesso. Depois, a autora realizou muitas alterações em seu texto, não só dos manuscritos para a primeira edição, mas também de uma edição para outra (ACKEL; MÓDOLO, 2019), o que torna o material suficientemente desafiador para um estudo paleográfico, ainda que em estágio inicial.

Além disso, o documento utilizado como teste foi produzido em uma folha de papel amarelo, cujo texto é datilografado com tinta preta e as rasuras, Lygia as fez com tinta azul. Esse fato facilita

bastante a edição de uma imagem, pois programas computacionais, como o *Sketchbook*, possuem ferramentas de rápida identificação e modificação de cores.

Imagens digitais são fáceis de manipular e editar. Pode-se adicionar ou remover características importantes de uma imagem sem deixar vestígios óbvios de adulteração. À medida que câmeras digitais e câmeras de vídeo substituem suas contrapartes analógicas, a necessidade de autenticar imagens digitais, validar seu conteúdo e detectar falsificações aumenta. Propostas para detecção de manipulação maliciosa com imagens digitais podem ser encontradas em diversos trabalhos como os de Fridrich, Soukal, Lukáš (2003) e Birajdar e Mankar (2013).

Uma das primeiras identificações de adulteração documental de que se tem notícia remonta ao século IX, tempo de Carlos, o Calvo, pela obra *Decretos de Pseudo-Isidoro*, que dispunha, dentre outras, de coleções sobre legislações de monarcas (*Capitularia Benedicti Levitae*), procedimentos criminais (*Capitula Angilramni*) e cartas com assinaturas dos papas dos primeiros três séculos.

Por obras recriadas como essas e tantas outras ao longo da história, todas elas revestidas de formas múltiplas e complexas, é que se procurou estabelecer métodos e regras para a verificação da autenticidade e integridade de um documento.

E hoje, doze séculos depois, com a prática ilícita ainda em vigor e cada vez mais aprimorada, a Paleografia reinventa-se com uma linguagem digital e recupera uma de suas premissas básicas que é a de fazer uma leitura correta de um documento e de atestar sua validade.

Nesse sentido, faz-se um convite aos pesquisadores de áreas afins para que conheçam os variados projetos em Paleografia Digital que têm sido implementados em universidades, por exemplo: (i) o pioneiro *Digital Resource and Database of Palaeography, Manuscripts and Diplomatic* (DigiPal), do King's College de Londres, que oferece imagens digitais de caligrafia medieval com descrições e caracterizações detalhadas da escrita, além de incorporar diferentes formas de explorar e manipular a informação, como imagens marcadas, mapas dinâmicos e linhas do tempo e processamento de imagens (KCL, 2020); (ii) o projeto *Cuneiform Digital Palaeography*, da Universidade de Birmingham e do Museu Britânico, custeado pela bolsa Leverhulme Trust, que procura estabelecer uma análise paleográfica bastante detalhada da escrita cuneiforme. O eixo central desse projeto é a construção de um banco de dados *online* que oferece imagens digitais de signos cuneiformes individuais retirados diretamente das fontes originais (UB, 2018); (iii) o projeto *Mulheres na América Portuguesa*, que tem como objetivo central sistematizar e dar acesso a um catálogo de documentos escritos por e sobre mulheres na América Portuguesa (NEHiLP, 2020).

1. Paleografia Digital

Ao longo dos anos, e particularmente na última década, cientistas da computação têm colaborado com estudiosos das ciências sociais, artes e humanas para construir ferramentas que facilitem seu trabalho.

O início dessa união se deu em 1949, quando Roberto Busa, um jesuíta e teólogo italiano, abordou Thomas J. Watson, fundador da IBM, buscando ajuda na indexação das obras de Tomás de Aquino (HOCKEY, 2004).

A criação de um índice que permitia o armazenamento e a recuperação de dados documentais de Aquino foi tão bem-sucedida que os mundos das humanidades e da computação estruturaram teorias e métodos transdisciplinares que formaram o que hoje conhecemos por *Humanidades Digitais*:

O campo das humanidades digitais não é unificado, mas uma série de práticas convergentes que exploram um universo em que: a) Impressão não é mais um meio exclusivo ou normativo para produzir ou disseminar conhecimentos, em vez disso, a impressão encontra-se absorvida por novas configurações multimídia; e b) ferramentas digitais, técnicas e mídia alteraram a produção e disseminação do conhecimento nas artes, ciências humanas e sociais. As humanidades digitais buscam desempenhar um papel inaugural em relação a um mundo em que, as universidades, que antes atuavam individualmente como produtores, administradores e disseminadores do conhecimento ou cultura, agora são chamadas a transformar seu discurso acadêmico em modelos digitais, para esferas públicas recém emergentes da presente era: o www, a blogosfera, as bibliotecas digitais etc.), para modelar a excelência e a inovação nesses domínios e para facilitar a formação de redes de produção, troca e disseminação que são, ao mesmo tempo, globais e locais.⁵

A intersecção do digital com as humanidades criou um ambiente no qual estas últimas se tornaram sujeitas a novas abordagens que levantaram questões sobre a natureza das humanidades, ao mesmo tempo que abriram novas teorias e métodos de pesquisa.

Por exemplo, a citação acima diz que “[...] a impressão encontra-se absorvida por novas configurações multimídia [...]”. A partir desse fato, a Paleografia que, por sua natureza, também se

⁵ “Digital Humanities is not a unified field but an array of convergent practices that explore a universe in which: a) print is no longer the exclusive or the normative medium in which knowledge is produced and/or disseminated; instead, print finds itself absorbed into new, multimedia configurations; and b) digital tools, techniques, and media have altered the production and dissemination of knowledge in the arts, human and social sciences. The Digital Humanities seeks to play an inaugural role with respect to a world in which, no longer the sole producers, stewards, and disseminators of knowledge or culture, universities are called upon to shape natively digital models of scholarly discourse for the newly emergent public spheres of the present era (the www, the blogosphere, digital libraries etc.), to model excellence and innovation in these domains, and to facilitate the formation of networks of knowledge production, exchange, and dissemination that are, at once, global and local”. (SCHNAPP et al., 2009, p. 01, tradução nossa).

ocupa do instrumento, tinta e, às vezes, do suporte⁶ (objeto de estudo da Codicologia), tem repensado sobre significações e resultados de análises em variados elementos, como assinaturas, carimbos, marcas d'água que, muitas vezes, a depender da resolução de um fac-símile, podem ser prejudicadas.

A variedade de plataformas, aplicativos, técnicas e ferramentas, todos desenvolvidos sob a rubrica de “digital”, tem mudado drasticamente a maneira como os humanistas trabalham, como pesquisam, reúnem informações, organizam, analisam e interpretam e divulgam as descobertas.

Uma reflexão que deve ser feita para garantir o avanço dessa transdisciplinaridade e, portanto, da produção e disseminação do conhecimento, é sobre a medida a que as ciências envolvidas devem se adaptar para obter seus resultados, de forma que não sejam minadas pelo automático, mas beneficiadas pela inclusão de novas metodologias.

Isto significa dizer que, ao serem agregados novos saberes a uma ciência, por exemplo, à Paleografia, deve-se pensar sobre quanto de conhecimento entre as disciplinas é necessário ter. Quanto de conhecimento paleográfico é necessário para elaborar uma pesquisa com recursos digitais, para a consulta de um manuscrito na internet, para a construção de um banco de dados, para a elaboração de um programa que leia manuscritos automaticamente, para avaliação de um tipo de escrita etc.? Ainda, quanto o cientista computacional precisará ter de conhecimentos paleográficos para elaborar algoritmos que resultem em análises decorrentes de leitura e escrita em manuscritos?

Aumentando o escopo dessas perguntas, Burdick *et al.* (2012, p. 13) fornecem uma definição mais ampla e inclusiva:

[As Humanidades Digitais] questionam o que significa ser um ser humano na era da informação em rede, que participa na prática de comunidades fluidas, perguntando e respondendo pesquisas que não podem ser reduzidas a um único gênero, meio, disciplina ou instituição... Trata-se de uma abordagem global, trans-histórica e transmidiática para o conhecimento e construção de significado.⁷

⁶ Refere-se aqui à disposição textual no suporte. No entanto, o termo *suporte* também sugere questionamentos que não serão detalhados neste artigo, mas podem ser observados, por exemplo, em um processador de texto ou de imagens. Como definir suporte digital? Pela tela do dispositivo, pelo programa que apresenta a folha digital ou pela própria folha digital? Ou ainda, só será suporte quando houver a impressão em papel físico? Este último argumento, se verdadeiro, contradiria o item *a* da citação acima. Questionamentos sobre o instrumento utilizado para a escrita, também no mesmo ambiente, podem ser feitos da mesma maneira.

⁷ Tradução nossa de “...asks what it means to be a human being in the networked information age and to participate in fluid communities of practice, asking and answering research questions that cannot be reduced to a single genre, medium, discipline, or institution... It is a global, trans- historical, and transmedia approach to knowledge and meaning-making.” (BURDICK *et al.*, 2012, p. 13, tradução nossa)

É claro que o primeiro fato a se notar é que pesquisas, hoje, em sua grande maioria, começam *online*.⁸ No entanto, não se pretende substituir o homem pela máquina. A observação humana, o contato físico entre o homem e seu objeto de estudo levam a questionamentos maiores e mais profundos do que uma análise automática. Deve-se integrar conhecimentos computacionais às atividades humanas, de modo que sua prática e ensino sejam aumentadas exponencialmente nas redes e comunidades acadêmicas, virtuais ou não.

De fato, a questão de saber se a computação em ciências humanas é uma disciplina humanística, uma ciência social ou uma disciplina científica/tecnológica ainda paira. Problemas que se relacionam com acesso pessoal ao objeto histórico também existirão. Além disso, dada a natureza intrinsecamente complexa das ciências em geral, parece improvável que métodos digitais possam produzir resultados “perfeitos” para muitas perguntas em um futuro próximo. (STOKES, 2015)

A proposta desse artigo não se restringe apenas a tais reflexões, mas também a mostrar possíveis meios de se aproveitar da computação para facilitar, melhorar, expandir e talvez até mesmo mudar a maneira do fazer paleográfico. Para isso, buscou-se apoio na Paleografia Digital (PD), que pode ser definida como a convergência entre os campos da Paleografia e a Computação, proposta, entre outros, por Ciula (2005) e Buzzetti (2011).

Considerando-se o estado da arte, é possível perceber que muito tem sido produzido e questionado no campo da PD, por exemplo, o trabalho da equipe de Stokes (2015) para encontrar junções em fragmentos do projeto *Cairo Genizah*. O cópuz é formado por milhares de fragmentos, ou seja, é demasiado grande para que o ser humano gerencie sozinho, mas, como testado em Wolf *et al* (2011a e b), o computador pode fazer sugestões de onde procurar.

Mesmo que apenas 1% dessas sugestões sejam precisas, isso ainda equivale a cerca de 3.500 novas junções, que não haviam sido encontradas antes. Em contraste com isso, está a verificação do pesquisador, em que os métodos podem ser os mesmos, mas a questão subjacente é significativamente diferente, pois não se buscam fragmentos semelhantes, mas, saber quais deles foram escritos pela mesma pessoa (STOKES, 2015). A diferença é como verificar o resultado.

Segundo Stokes (2015), ainda não está claro como saber se a semelhança identificada pelo computador necessariamente significa produção pelo mesmo escriba, particularmente para casos em que não há nenhuma verdade fundamental para testar os métodos. Assumindo que não existe um

⁸ Felizmente, ainda há muitos pesquisadores que adentram arquivos e bibliotecas para estudar seus objetos de estudo. O acesso pessoal à documentação leva a questionamentos sobre análises que, muitas vezes, são mais profundas quando trabalhadas diretamente nos originais, por exemplo, segurando o suporte, manuseando-o; observando sua consistência, qualidade, textura; medindo sua dimensão, formato, superfície escrita, espaçamento entre linhas; encontrando rasgos, deteriorações; identificando técnicas utilizadas na escrita, avaliando a pressão que o escriba fez com seu instrumento sobre o suporte etc.

algoritmo perfeitamente confiável, qualquer resultado será fornecido na forma de probabilidade e caberá ao paleógrafo tomar a decisão final.

Volta-se à questão da atividade humana em contraste com a automática e é relevante notar que Derolez (2003a) já havia pensado nesses problemas quando questionou meios para provar uma descrição gráfica.

O aspecto diacrônico dos estudos da PD parece ser também um desafio, uma vez que os métodos aplicados para desenvolvimento da disciplina tendem a se concentrar na classificação ao invés da evolução da escrita (STANSBURY, 2009).

Além dos projetos já citados anteriormente, podem ser destacados os trabalhos de Padma e Vijaya (2009), sobre a identificação de escritas em Tegulu, Devanagari e Inglês por meio da comparação; de Wolf *et al.* (2011), sobre ferramentas digitais para reconhecimento de caracteres em manuscritos históricos; de Kestemont e Stutzmann (2017), sobre a identificação de manuscritos latinos medievais, por meio de redes convolucionais neurais.⁹

Algum trabalho foi feito na datação de manuscritos com base na suposição de que variações na escrita se desenvolvem ao longo do tempo (HE *et al.*, 2014), mas o objetivo ainda de determinar datas, em vez de investigar questões como a evolução em si não pôde ser alcançado (cf. POULLE, 1977).

As questões levantadas aqui não são novas e, como observado nas referências, muitas já foram consideradas antes de se falar em Humanidades Digitais. As apostas são altas, na medida em que respostas eficazes a essas perguntas teriam implicações muito amplas, não apenas para os estudos paleográficos, mas para outras ciências humanas.

Para além de questões metodológicas de reconhecimento ótico de caracteres,¹⁰ proposta já bastante consolidada na literatura (GUPTA, JACOBSON, GARCIA, 2006; CHOUDHAR, RISHI, SAVITA, 2013; PEKETI, MAHESH, UMA, VISHISTHA, 2021), ao se pensar no termo *Paleografia Digital*, deve-se conceber, em primeiro lugar, que um computador relaciona valores que lhe são fornecidos por meio de dados e, quando solicitado a processá-los, faz uma combinação para obter um resultado. A exemplo, tomem-se uma calculadora, com operações matemáticas; uma câmera de um celular, com configurações de luminosidade e foco; um *site* de pesquisa com filtros para encontrar um manuscrito específico.

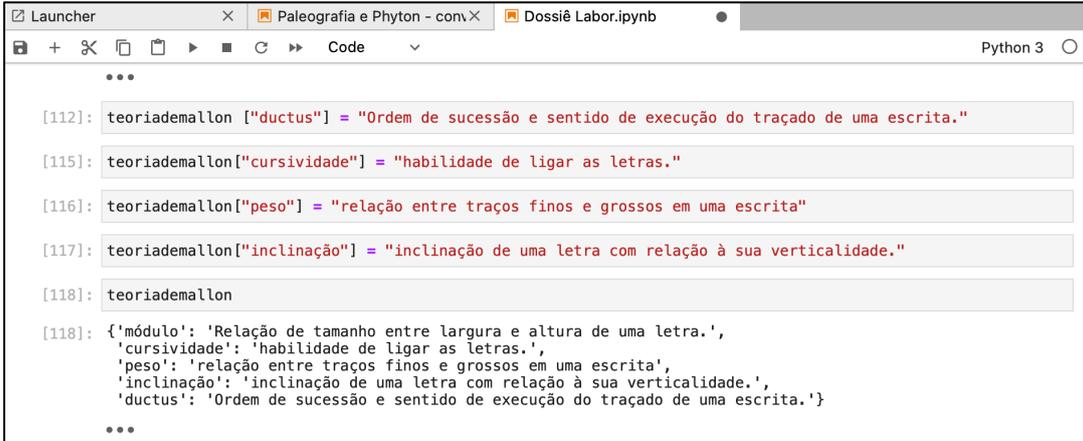
⁹ Trata-se de uma arquitetura de redes especialmente desenhada para classificar imagens por meio de um cálculo de camadas. (ACKEL e MADEIRA, 2021)

¹⁰ Tradução nossa de *optical character recognition*, que se refere à habilidade da máquina em utilizar sua câmera para ler um texto (impresso ou manuscrito) e convertê-lo em um documento transcrito e editável, como feito com um telefone celular em aplicativos de tradução, por exemplo.

Numa proposta de transposição de conhecimentos computacionais para os da paleografia, propõe-se um conceito menos abstrato quanto a regras e terminologias da computação. Assim, abaixo apresentam-se dois exemplos do tipo de informação que a máquina precisa para relacionar, processar e oferecer resultados.

A figura 2 mostra algoritmos – fórmulas, receitas, instruções –, que são digitados em uma programação de linguagem *Python*¹¹. Nesse caso, a “máquina está aprendendo”¹², de acordo com Mallon (1952), quais características podem ser identificadas em um tipo de escrita.

Figura 2 – Algoritmos para definição de termos da teoria do desenvolvimento da escrita de Mallon (1952)



```

[112]: teoriademallon ["ductus"] = "Ordem de sucessão e sentido de execução do traçado de uma escrita."
[115]: teoriademallon["cursividade"] = "habilidade de ligar as letras."
[116]: teoriademallon["peso"] = "relação entre traços finos e grossos em uma escrita"
[117]: teoriademallon["inclinação"] = "inclinação de uma letra com relação à sua verticalidade."
[118]: teoriademallon
[118]: {'módulo': 'Relação de tamanho entre largura e altura de uma letra.',
      'cursividade': 'habilidade de ligar as letras.',
      'peso': 'relação entre traços finos e grossos em uma escrita',
      'inclinação': 'inclinação de uma letra com relação à sua verticalidade.',
      'ductus': 'Ordem de sucessão e sentido de execução do traçado de uma escrita.'}

```

Fonte: elaboração própria.

A partir da ilustração acima, pode-se compreender a funcionalidade de um banco de dados em que todas as palavras (valores), servirão, não só de filtros de busca na procura do significado de termos, mas também de resultados que a máquina poderá fornecer, por meio de uma imagem encontrada em algum manuscrito. Esses procedimentos iniciais já abrem amplo horizonte de agrupamento de informações que o paleógrafo poderá construir para analisar seu objeto textual.

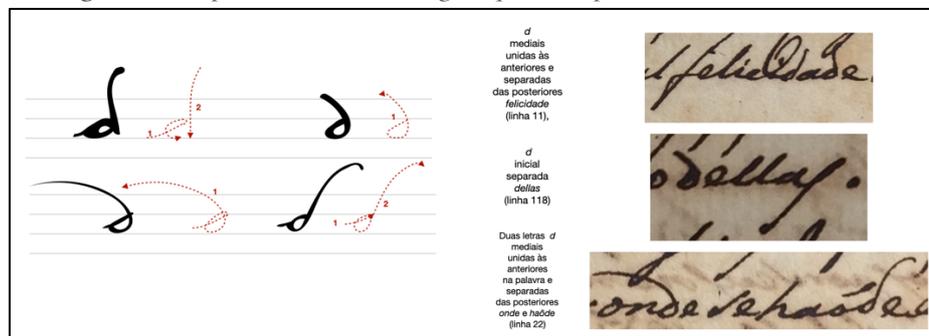
No segundo exemplo, na imagem abaixo, um outro tipo de associação é ilustrado. A máquina compreende o texto como imagem e significado, quer dizer, a escrita manual é concebida por meio de duas variáveis: o texto como forma (parâmetros morfológicos) e o texto como significado (parâmetros semânticos), que, no caso da figura 3 mostra uma letra como um conjunto de traços que formam imagens com significação, neste caso, variadas formas da letra *d*. As imagens foram retiradas

¹¹Essa linguagem foi selecionada como exemplo por sua atualidade, facilidade de uso, grande disponibilidade de exemplos e algoritmos prontos em *sites* da internet e por dialogar com outras linguagens, como C e C++.

¹²O termo se refere ao inglês *machine learning*, a habilidade de uma máquina de exercer determinadas funções sem supervisão humana. Para conceitos elementares, veja-se Herbrich e Graepel (2003).

de um manuscrito datado em 1650, cuja autoria é atribuída ao Padre Antonio Vieira (ACKEL, 2019).¹³

Figura 3 – Proposta de *ductus* e a imagem que corresponde à letra *d* no manuscrito



Fonte: elaboração própria.

Para a realidade computacional, cada forma de letra deve ser dividida por seus traços e cada traço deve ser representado por sua direção e sentido. Assim também é com o homem, no ensino da escrita: ele aprende com seu próprio movimento a copiar as sequências de uma letra, e a ela, atribuir-lhe o valor que é dado.

Ainda que, até hoje, o primeiro estudioso moderno a aplicar o termo *ductus* à escrita tenha sido Bernard de Montfaucon (1708), a Escola Paleográfica Francesa, por meio de Jean Mallon (1952), reintroduziu o princípio do *ductus* nos estudos modernos e demonstrou sua influência norteadora em todas as mudanças da forma das letras que ocorreram na história da escrita romana.

Seu trabalho, e o de seus seguidores, continua a ser instrumental no desenvolvimento da PD, como pode ser observado no exemplo acima e em outros trabalhos como os de Guinet e Kandel (2010) e Simonnet, Anquetil, Bouillon (2017).

Conforme os estudos fundamentais de Ciulla (2005), dar forma à aparência gráfica manuscrita de um lado (conceber o texto como imagem), e processar os dados associados ao texto como porção linguística (texto como significado), do outro, mostra o horizonte de análises que se abre para o campo da PD.

Dentre as múltiplas sistematizações possíveis, a partir dessas duas variáveis, poder-se-ia criar hipóteses sobre a evolução morfológica ao longo do tempo, sobre a existência ou não de padrões de ciclos escriturários, ou ainda, sobre ferramentas que identificariam modificações nas mãos do escriba.

¹³ O manuscrito foi utilizado como exemplo por fazer parte da constituição do estudo de alfabetos já citado. Está disponível para consulta sob o código de referência AL 135-015 e armazenado na caixa 18, códice 135, sala 01, da coleção Alberto Lamego, do Instituto de Estudos Brasileiros da Universidade de São Paulo.

É fato que, com o aumento da variedade e necessidade de ferramentas digitais, a Paleografia caminha para constituir-se uma disciplina inovadora na pesquisa e no ensino. Resta, assim, explorar possibilidades de descrições alternativas e complementares.

Portanto, na seção seguinte, apresenta-se uma nova hipótese para avaliação qualitativa com uso de ferramentas digitais. O trabalho se dá pela supressão de elementos que não fazem parte do conteúdo textual proposto na obra do autor. Neste caso, um manuscrito datilografado e rasurado de *Seminário dos Ratos*, de Lygia Fagundes Telles.

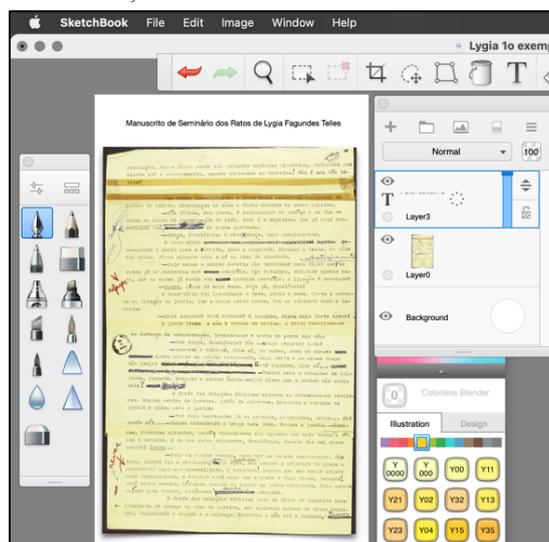
2. Proposta de utilização do *Sketchbook* para apagamento de rasuras

Visando ganho de velocidade e rigor com o auxílio do processamento de dados na utilização de métodos tradicionais paleográficos para análise visual de formas gráficas, propõe-se um procedimento específico que foi desenvolvido para o presente estudo. O objetivo é permitir o registro e a exploração de dados métricos, baseados em etapas que envolvem a adição de novas camadas e a remoção cuidadosa de pontos de obliteração, até que o texto original se torne visível usando uma ferramenta de pincel que é definida como a cor de fundo.

O programa computacional utilizado chama-se *Sketchbook* e foi desenvolvido para pintura e desenho à mão, para permitir que projetos nas áreas das artes gráficas sejam executados.

O primeiro passo foi a inclusão do fac-símile no programa, como mostrado na figura abaixo. Note-se a quantidade de intervenções que a autora fez, inclusive anotações laterais.

Figura 4 – Inserção do documento a ser trabalho no *Sketchbook*

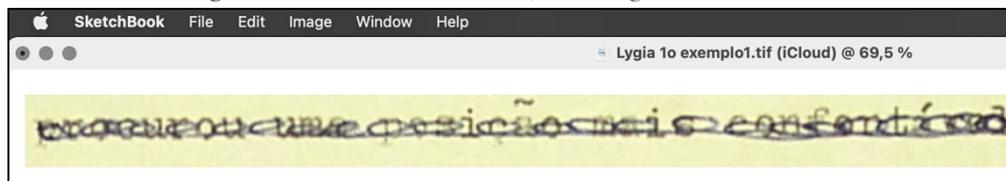


Fonte: elaboração própria.

Na sequência, isola-se o fragmento a ser editado, na primeira camada do arquivo.

A cada inserção de camadas, uma nova edição é feita.

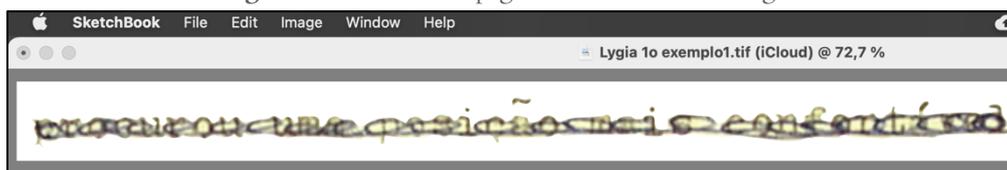
Figura 5 – Camada 1 com a inserção do fragmento a ser editado



Fonte: elaboração própria.

Após o isolamento da imagem, o recurso de eliminação de fundo é aplicado, dessa forma, percebe-se que o texto começa a se destacar, especialmente com relação aos diacríticos. Trata-se de um trabalho manual, neste caso, com a utilização de uma mesa digitalizadora da marca *Wacom*, modelo *CTL 4100*¹⁴. O uso desse tipo de instrumentos possibilita melhor precisão na aplicação dos filtros que o programa disponibiliza.

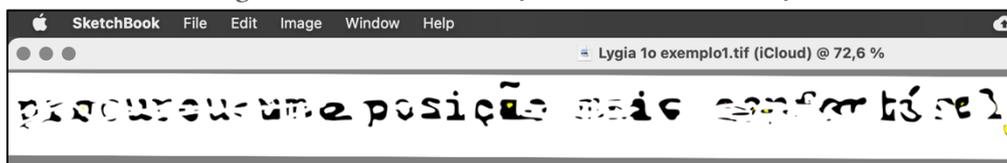
Figura 6 – Camada 2 - apagamento do fundo da imagem



Fonte: elaboração própria.

Em seguida, os valores de contraste e iluminação são modificados de forma que realcem o conteúdo datilografado e a ferramenta de substituição de cor é também aplicada, quer dizer, apaga-se somente a cor azul que a autora utilizou para rasurar o texto. Nessa etapa o texto, que antes estava por baixo da rasura, toma forma e permite conjecturação ou decifração.

Figura 7 – Camada 3 – saturação de contraste e iluminação

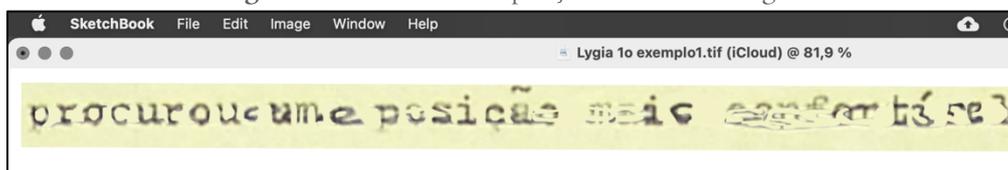


Fonte: elaboração própria.

¹⁴ Para detalhes sobre o uso de instrumentos e suportes digitais, cf. Ackel e Módolo (2020d).

A última etapa do processo devolve à imagem seu fundo, que neste caso é a cor amarela, original do papel utilizado para verificar se alguma outra parte do fragmento possa ser destacada, assim, com menos dificuldades, pode-se ler “*procurou uma posição mais confortável*”.

Figura 8 – Camada 4 – recuperação do fundo da imagem

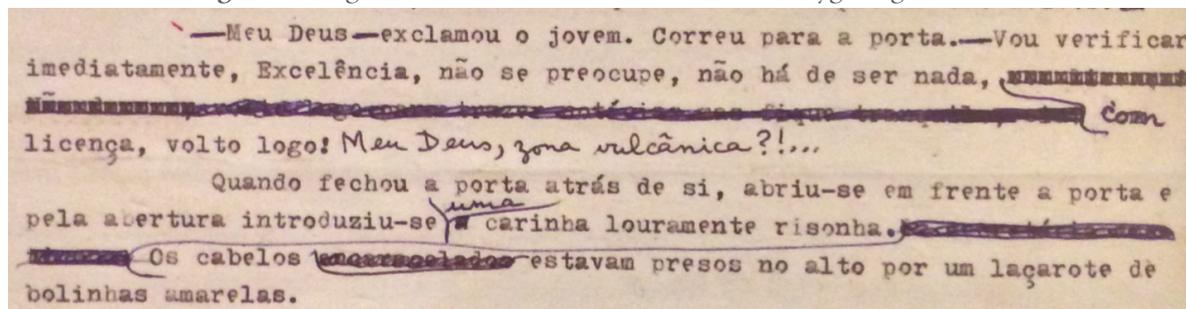


Fonte: elaboração própria.

O método de reconstituição de escrituras sob rasuras é baseado em funções semelhantes a "Reflectance Transformation Imaging" e "Virtual Vellum" da CHI, embora a aqui apresentada mostre aplicações específicas para a recuperação de texto rasurado.

A maior parte do trabalho que tem sido desenvolvido, até agora, tem se concentrado em rasuras, como no caso dos manuscritos apresentados. Um desafio é, portanto, garantir que toda a diversidade de rasura sobre a escrita seja retirada. Como exemplo, mostra-se a figura abaixo, que tem se mostrado especialmente desafiadora nessa questão.

Figura 9 – Fragmento do manuscrito IMS/LFT-08.014, de Lygia Fagundes Telles



Fonte: Instituto Moreira Salles (2018).

Considerações finais

O estudo realizado nas intervenções autorais de um documento, neste caso, rasuras para correção do texto editadas em um programa computacional, permitiu enxergar um horizonte de possibilidades na prática paleográfica adicionada ao campo das Humanidades Digitais. O espectro de descobertas justifica-se por sua pertinência e importância no tocante à constituição de práticas acadêmicas que valorizem as competências de leitura paleográfica.

O sistema de edição de imagem que incorpora técnicas inteligentes e semiautomáticas para melhorar a eficiência do processo de edição é aplicado em imagens individuais. As funções desempenhadas pelo sistema são capazes de realizar operações que removem objetos do conjunto de dados. Além de rastreios automáticos, edições manuais também foram executadas para permitir operações gerais de edição. Essas técnicas, aliadas a uma interface de usuário final, contribuem para a melhoria da velocidade do processo de edição quando aplicadas a casos reais.

No exemplo apresentado, o programa destinou-se à filtragem preliminar de imagens obtidas a partir do documento selecionado, por meio de métodos ópticos para a eliminação de elementos indesejáveis na imagem – rasuras –, como transformações de pontos de iluminação, aumento de contraste e nitidez.

O resultado é uma leitura, ainda conjecturada, mas facilitada, do que a autora teria escrito em concepções textuais anteriores. Este tipo de atividade paleográfica apresenta também a vantagem de permitir acompanhar com clareza as versões por que foi passando o texto, tornando mais fácil e intuitiva a compreensão do processo da sua composição. Tal fato, leva a pesquisa a outros campos do saber, que mostram que Paleografia não é apenas a ciência que estuda as escrituras antigas, mas uma disciplina renovada que considera o estudo global da história dos usos e práticas da escrita.

Embora este texto tenha se concentrado em um estudo de caso específico, a metodologia aqui explorada reivindica uma abrangência mais ampla. O que os resultados mostraram de fato é que, cada edição feita em nova camada permite uma descrição analítica de pequenos detalhes modificados no texto para comparação sistemática com outros documentos. Como consequência, fica claro que a natureza do método paleográfico tradicional é enriquecida em vez de diminuída ou subestimada.

Trata-se de uma observação cuidadosa, especialmente nesta era digital que velozmente propaga informações. Porém, o ato de se analisar com os olhos humanos um documento histórico não pode deixar de constituir-se, em detrimento de novas tecnologias. A brevidade do tempo dedicado a pesquisas acadêmicas em acervos virtuais bibliográficos parece estar se acelerando cada vez mais, diante da segurança de preservação de obras sempre disponíveis, nos sites, nas nuvens, para consulta, inclusive aquelas raras, julgadas perdidas.

Neste artigo, foram apresentadas possibilidades de análises metodológicas no que tange à paleografia digital. Com a intenção de aprofundar o conhecimento sobre os elementos gráficos de testemunhos, a metodologia explorada, acredita-se, pode ser aplicada não apenas a outros estudos de caso e *corp*us, mas também aprimorar projetos existentes em edições digitais baseadas em imagens de manuscritos, oferecendo algumas ferramentas eficazes para a análise paleográfica.

Além de tentar melhorar a precisão dos resultados, então, também vale a pena perguntar o que pode ser feito de forma útil agora com resultados que oferecem análises qualitativas.

Nesse sentido, acredita-se que manter e expandir a paleografia digital não apenas estimulará o interesse de pesquisadores a se engajarem ativamente nessa abordagem, como também promoverá esforços de recuperação de fatos históricos, contidos em documentos ainda anônimos em bibliotecas e arquivos, o que irá alimentar a próxima geração de acadêmicos.

Referências bibliográficas

- ACKEL, A. Estudo paleográfico de manuscrito do século XVII. **Revista Todas as Letras**. São Paulo, v. 21, p. 1-23, 2019. Disponível em: <<http://editorarevistas.mackenzie.br/index.php/tl/article/view/12308>>. Acesso em: 22 out 2020.
- ACKEL, A. Estudos paleográficos: gênese, evolução e tendências. **Revista Todas as Letras**. São Paulo, v. 23, n. 1, p. 1-26, 2021. Disponível em: <<http://editorarevistas.mackenzie.br/index.php/tl/article/view/14367>>. Acesso em: 9 nov 2021.
- ACKEL, A.; MÓDOLO, M. Leitura e escrita de manuscritos: pena e suporte digitais na prática paleográfica. **Filologia e Linguística Portuguesa**. São Paulo, v. 22, n. 2, p. 233-245, jul./dez. 2020. Disponível em: <<https://www.revistas.usp.br/flp/article/view/176142>>. Acesso em: 15 fev 2021.
- ACKEL, A.; MADEIRA, F. Os caminhos da codicologia. **Revista Todas as Letras**. São Paulo, v. 23, n. 1, p. 1-15, 2021. Disponível em: <<http://editorarevistas.mackenzie.br/index.php/tl/article/view/14359>>. Acesso em: 09 nov 2021.
- BIRAJDAR, G.; MANKAR, V. Digital image forgery using passive techniques: A survey. **Digital Investigation**. v. 10, n. 3, 2013, p. 226-245. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1742287613000364>>. Acesso em: 10 set 2020.
- BOND, Edward; THOMPSON, Maunde; WARNER, George. **The Palæographical Society**. Facsimiles of manuscripts and inscriptions. London: [s.n.], 1873. Disponível em: <https://archive.org/details/ldpd_10836923_005/page/n1/mode/2up>. Acesso em: 10 set 2020.
- BURDICK, A.; J. DRUCKER, LUNENFELD, P.; PRESNER, T.; SCHNAPP. **Digital Humanities**. Cambridge, Massachusetts: MIT Press, 2012. Disponível em: <<https://mitpress.mit.edu/books/digitalhumanities>>. Acesso em: 3 out 2021.
- BUZZETTI, D. Digital Text Representation: Expression and Content. In FORD, D. **Contexts: Proceedings of ANPA 31, (Alternative Natural Philosophy Association, 31st International Meeting, Wesley House, Cambridge, August 2010)**. London: ANPA, 2011, pp. 124-145. Disponível em: <https://www.academia.edu/1246494/Digital_Text_Representation_Expression_and_Content?email_work_card=view-paper>. Acesso em: 10 set 2020.
- CHABRIES, Douglas; BOORAS, Steven; BEARMAN, Gregory. Imaging the Past: Recent Applications of Multispectral Imaging Technology to Deciphering Manuscripts. **Antiquity**. v. 77, n. 296, 2013, p. 359-72. Disponível em: <<http://www.digitalhumanities.org/dhq/vol/8/4/000187/000187.html>>. Acesso em: 10 set 2020.
- CHOUDHARY, A.; RISHI, R.; SAVITA, A. A New Character Segmentation Approach for Off-Line Cursive Handwritten Words. **Procedia Computer Science**. v. 17, 2013, p. 88-95.

Disponível em: <<https://www.sciencedirect.com/science/article/pii/S1877050913001464>>.

Acesso em: 22 out 2020.

CIULLA, A. Digital palaeography: using the digital representation of medieval script to support palaeographic analysis. **Digital Medievalist**. v.1, 2005. Disponível em: <<https://journal.digitalmedievalist.org/articles/10.16995/dm.4/>>. Acesso em: 10 set 2020.

CIULLA, A. Digital Palaeography: what is it about? **Digital Scholarship in the Humanities**. v. 32, n. 2, 2017, p. 115- 123. Disponível em: <https://academic.oup.com/dsh/article/32/suppl_2/ii89/4259068>. Acesso em: 10 set 2020.

DEROLEZ, A. **Painted labyrinth**: the world of the Lindisfarne Gospels. London: The British Library, 2003.

EDWARDS, N.; WEBB, S.; KREST, C.; CAMPEN, Douglas.; MANNING, P.; BERGMAND, U.; WOGELIUS, A. A new synchrotron rapid-scanning X-ray fluorescence (SRS-XRF) imaging station at SSRL beamline 6-2. **Journal of Synchrotron Radiation**. v. 25, part 5, set 2018, p. 1565-1573. Disponível em: <<https://journals.iucr.org/s/issues/2018/05/00/il5009/>>. Acesso em: 10 set 2020.

FRAUZAURIH. **Ductus**: la formation de l'alphabet moderne. Paris: Frauzaurih, 2016. 1 vídeo (21:33 minutos). Disponível em: <<https://www.youtube.com/watch?v=bMBmNBUhMW0>>. Acesso em: 09 nov 2021.

FRIDRICH, J.; SOUKAL, D.; LUKÁŠ, J. Detection of Copy-Move Forgery in Digital Images. **CiteSeerX**. Pennsylvania State University, 2003, p. 213-222. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.121.1962>>. Acesso em: 10 set 2020.

GUINET, E.; KANDEL, S. Ductus: A software package for the study of handwriting production. **Behaviour Research Methods**. v. 46, 2010, p. 326-332. Disponível em: <<https://link.springer.com/article/10.3758/BRM.42.1.326>>. Acesso em: 28 out 2020.

GILISSEN, L. **Prolégomènes à la codicologie**: recherches sur la construction des cahiers et la mise en page des manuscrits médiévaux. Gand: Story-Scientia, 1977. Disponível em: <https://www.persee.fr/doc/bec_0373-6237_1979_num_137_1_450158_t1_0084_0000_2>. Acesso em: 21 jan 2021.

GUPTA, M.; JACOBSON N.; GARCIA, E. OCR binarization and image pre-processing for searching historical documents. **Pattern Recognition**. v. 40, 2007, p. 389-397. Disponível em: <<https://www.sciencedirect.com/science/article/abs/pii/S0031320306002202>>. Acesso em: 22 out 2020.

HAVENS, H. Adobe Photoshop and Eighteen-Century Manuscripts: A New Approach to Digital Paleography. **Digital humanities Quarterly**. v. 8, n4, 2014. Disponível em: <<http://www.digitalhumanities.org/dhq/vol/8/4/000187/000187.html#d71243e206>>. Acesso em: 10 set 2020.

- HE, S.; SAMARA, P.; BURGERS, J.; SCHOMAKER, J. Towards Style-Based Dating of Historical Documents. Comunicação apresentada na **International Conference on Handwriting Recognition**. 2014, Creta. Disponível em: <http://www.researchgate.net/publication/263518853_Towards_stylebased_dating_of_historical_documents>. Acesso em: 15 out 2020.
- HERBRICH, R.; T. Introduction do the Special Issue on Learning Theory. **Journal of Learning Machine**. v. 4, 2003, p. 755-757. Disponível em: <<https://www.jmlr.org/papers/volume4/herbrich03a/herbrich03a.pdf>>. Acesso em: 10 set 2020.
- HOCKEY, Susan. The History of Humanities Computing. **A companion to digital humanities**. Oxford: Blackwel. 2004. Disponível em: <<http://www.digitalhumanities.org/companion/>>. Acesso em: 21 set 2020.
- IMS - Instituto Moreira Salles. **IMS**. Rio de Janeiro, RJ: 2018. Disponível em: <<https://ims.com.br/titular-colecao/lygia-fagundes-telles/>>. Acesso em: 09 nov 2021.
- KCL - King's College London. **Digital Resource and Database of Palaeography, Manuscripts and Diplomatic: DigiPal**. London: 2020. Disponível em: <<https://kclpure.kcl.ac.uk>>. Acesso em: 09 nov 2021.
- KESTEMONT, M.; STUTZMANN, D. Script identification in medieval Latin manuscripts using convolutional neural networks. **Digital humanities**. Book of abstracts. Montreal, August 10, 2017, p. 283-285. Disponível em: <<https://www.semanticscholar.org/paper/Script-Identification-in-Medieval-Latin-Manuscripts-KestemontStutzmann/f8ce8dad0dccac8acaece2da9f2565b384f115ca>>. Acesso em: 15 out 2020.
- MABILLON, Jean. **Librorum de re diplomatica supplementum**: in quo archetypha in his libris pro regulis proposita, ipsaeque regulae denuo confirmantur, novisque speciminibus & argumentis asseruntur & illustrantur. Paris: Sumtibus Caroli Robuste, 1704. Disponível em: <https://archive.org/details/gri_33125008690535>. Acesso em: 10 set 2020.
- MALLON, J. **Paléographie romaine**. Madrid: Instituto Antonio de Nebrija de Filoogia, 1952. Disponível em: <[https://archive.org/search.php?query=mallon&and\[\]=mediatype%3A%22texts%22&and\[\]=year%3A%221955%22&and\[\]=year%3A%221949%22](https://archive.org/search.php?query=mallon&and[]=mediatype%3A%22texts%22&and[]=year%3A%221955%22&and[]=year%3A%221949%22)>. Acesso em: 22 aug 2020.
- MONTFAUCON, B. **Palæographia Græca sive de ortu et progressu literarum græcarum, et de variis omnium sæculorum Scriptionis Græcæ generibus**: itemque de Abbreviationibus & de Notis variarum Artium ac Disciplinarum. Paris: L. Guérin, J. Boudot e C. Robustel, 1708. Disponível em: <<http://gallica.bnf.fr/ark:/12148/bpt6k1041709q/f7.image>>. Acesso em: 22 out 2020.
- NEHiLP - Núcleo de Apoio à Pesquisas em Etimologia e História da Língua Portuguesa. **Mulheres na América Portuguesa - M.A.P.** São Paulo, 2020. Disponível em: <<http://map.prp.usp.br>>. Acesso em: 09 nov 2021.

- PADMA M.; VIJAYA, P. Identification of Telugu, Devanagari and English Scripts using **Discriminating Features**. *International Journal of Computer Science and Information*. v. 1, n. 2, p. 64–78, 2009. Disponível em: <<http://citeseerx.ist.psu.edu/viewdoc/summary?doi=10.1.1.212.7161>>. Acesso em: 15 out 2020.
- PALMA, G.; CORSINI, M.; CIGNONI, P.; SCOPIGNO, R.; MUDGE, M. Dynamic shading enhancement for reflectance transformation imaging. *ACM Journal on Computing and Cultural Heritage*. v. 3, n. 2, 2010, p. 60–82. Disponível em: <http://culturalheritageimaging.org/What_We_Do/Publications/acmdl2010/index.html>. Acesso em: 10 set 2020.
- PEKETI D.; MAHESH V.; UMA R.; VISHISTHA, M. Web based optical character recognition application using flask and tesseract. *Materials Today: Proceedings*, 2021, p. 784, 801. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S221478532038487X>>. Acesso em: 22 out 2020.
- POULLE, E. Une histoire de l'écriture. *Bibliothèque de l'École des Chartes*. v. 135, 1977, p. 137–44. Disponível em <https://www.persee.fr/doc/bec_0373-6237_1977_num_135_1_450098>. Acesso em: 20 out 2020.
- SCHNAPP, P.; LUNENFELD, P.; PRESNER, T. **The Digital Humanities Manifesto**. Engaged Digital Humanities Scholarship. Disponível em: <<http://humanitiesblast.com/publications/>>. Acesso em: 13 out 2020.
- SIMONNET, D.; ANQUETIL, E.; BOUILLON, M. Multi-criteria handwriting quality analysis with online fuzzy models. *Pattern Recognition*. v. 69, 2017, p. 310–324. Disponível em: <<https://www.sciencedirect.com/science/article/pii/S0031320317301474>>. Acesso em: 4 nov 2020.
- SKETCHBOOK - For everyone who loves to draw. **Sketchbook**. Los Angeles, CA: 202. Disponível em: <<https://www.sketchbook.com/apps>>. Acesso em: 09 nov 2021.
- SMITH, Albert. The photography of manuscripts. *London Medieval Studies*. v. 1, 1983, p. 179–207. Disponível em: <<https://www.cambridge.org/core/journals/antiquity/article/abs/imaging-the-past-recent-applications-of-multispectral-imaging-technology-to-deciphering-manuscripts/E6C39142EEA0C5C66A1E644D0D99873A>>. Acesso em: 10 set 2020.
- UB - University of Birmingham. **Cuneiform Digital Palaeography**. Birmingham, 2019. Disponível em: <<https://research.birmingham.ac.uk/>>. Acesso em: 09 nov 2021.
- WOLF, L.; POTIKHA, N.; DERSHOWITZ, R.; SHWEKA, Y.; CHOUKEA, A. Computerized paleography: Tools for historical manuscripts. **Proceedings: International Conference on Image Processing, ICIP**, 2011. Disponível em: <<https://doi.org/10.1109/ICIP.2011.6116481>>. Acesso em: 15 out 2020.