

Laboratório Virtual de
Humanidades Digitais

LaViHD

Humanidades Digitais | USP • LaPeLinC | UESB

Anteprojeto

para Acordo de Cooperação Acadêmica Nacional

Prof^ª. Dr^ª. Maria Clara Paixão de Sousa | USP
Prof^ª. Dr^ª. Cristiane Namiuti Temponi | UESB

1 Apresentação

O Laboratório Virtual de Humanidades Digitais (LaViHD) será um ambiente virtual para o desenvolvimento de ferramentas para acervos digitais e corpora eletrônicos, em particular para captura de imagens, edição filológica e anotação linguística. O LaViHD é um projeto conjunto da Universidade de São Paulo e da Universidade Estadual do Sudoeste da Bahia, e representa a consolidação de parcerias entre o Grupo de Pesquisas Humanidades Digitais da USP e o Laboratório de Pesquisas em Linguística de Corpus (LaPeLinC) da UESB, em torno da concepção e desenvolvimento de recursos digitais focalizados no estudo histórico da língua e na curadoria digital de acervos memoriais.

2 Objetivos

O objetivo amplo e de longo prazo do Projeto é consolidar o LaViHD como um pólo incubador e difusor de tecnologias de processamento de texto que, a um tempo, reflitam o estado da arte das tecnologias em âmbito internacional e sigam um compromisso rigoroso com as melhores práticas filológicas. O objetivo específico da proposta no horizonte dos próximos três anos será construir um ambiente digital colaborativo fechado (LaViHD^{beta}), dedicado ao desenvolvimento dos catálogos e das ferramentas utilizadas no processamento de textos dos projetos correntes em cada um dos grupos, visando avanços no desenvolvimento de ambos. Esses avanços de lado a lado serão possíveis graças a um intercâmbio de competências complementares, como justificado abaixo.

3 Justificativas

A proposta se justifica em dois planos. No plano dos objetivos de curto prazo, o intercâmbio de experiências entre os dois grupos de pesquisa e a transferência de tecnologias já desenvolvidas em cada um significaria um imenso avanço em cada Projeto, por conta da complementaridade de suas respectivas competências. Em particular, os objetivos tecnológicos do grupo da USP ganharão impulso graças ao caráter inovador dos recursos computacionais desenvolvidos pelo grupo da UESB para anotações e buscas linguísticas, e o grupo da UESB se enriquecerá graças à tradição do trabalho filológico que caracteriza o grupo da USP – tanto na perspectiva do saber filológico clássico, como em uma perspectiva tecnológica, como pólo originário do eDICTOR, o primeiro software para edições filológicas eletrônicas inteiramente desenvolvido no Brasil (Paixão de Sousa e Kepler, 2007; Paixão de Sousa, Kepler e Faria, 2010).

No plano dos objetivos de longo prazo, o desenvolvimento e a difusão das tecnologias de processamento voltadas ao tratamento de documentos históricos faz-se extremamente necessária frente ao adensamento de projetos que pretendem digitalizar e ou produzir edições filológicas em formato eletrônico, de um lado, e a falta de padrões consensuais na área, no caso das

pesquisas brasileiras. Isso dificulta imensamente o intercâmbio entre as pesquisas e impede, por vezes, a replicação e o reuso de recursos muitas vezes desenvolvidos com imenso investimento material e humano. A ideia da abertura do LaViHD para a comunidade de pesquisa justifica-se portanto na medida em que pretendemos colocar à disposição protocolos e ferramentas concebidos, fundamentalmente, pelo princípio de convergência. Importa ressaltar, nesse sentido, que partimos do compromisso com as tecnologias transferíveis e o acesso aberto, sendo nosso objetivo a difusão e democratização das tecnologias desenvolvidas e a desenvolver.

4 Produtos

Os principais produtos da proposta serão:

- O próprio LaViHD, ambiente virtual de desenvolvimento de corpora e ferramentas – em versão testes (LaViHD^{beta}) e em versão estável;
- O eDictor 2.0, uma ferramenta para edições filológicas eletrônicas inteiramente baseada na Web

Além disso, pretende-se preparar publicações em periódicos científicos com relatos e discussões sobre a experiência de desenvolvimento das ferramentas, e workshops sobre ferramentas de catalogação e anotação de corpus.

5 Histórico e contexto da parceria

Os objetivos e os interesses de pesquisa do LaPeLinC e do Grupo de Pesquisas Humanidades Digitais apresentam afinidades muito marcadas, tanto do ponto de vista temático como metodológico, que impulsionam a presente proposta.

O Grupo de Pesquisas Humanidades Digitais (<http://humanidadesdigitais.org>) atua desde 2009, na Universidade de São Paulo; entre 2009 e 2012, o grupo esteve sediado na Biblioteca Brasileira Guita e José Mindlin, como parte da equipe que concebeu, construiu e administrou a primeira plataforma da Biblioteca Brasileira Digital (2008-2013). Hoje, está abrigado junto ao Núcleo de Apoio à Pesquisa em Etimologia e História da Língua Portuguesa, NEHiLP. Desde 2017, grupo vem se dedicando centralmente ao Projeto M.A.P. (Mulheres na América Portuguesa): Mapeamento de escritos de mulheres e sobre mulheres no espaço atlântico português a partir de métodos das Humanidades Digitais (<http://map.prp.usp.br/>), sob coordenação da Profa. Dra. Maria Clara Paixão de Sousa e da Profa. Dra. Vanessa Martins do Monte, em diversas etapas (cf. Monte e Paixão de Sousa, 2017a; Paixão de Sousa, 2018; Monte, 2018; Paixão de Sousa, 2019). O M.A.P. tem como objetivo central sistematizar e tornar visível para pesquisas futuras um conjunto de fontes documentais imensamente importantes para os estudos filológicos e para os estudos da história da língua, da história social, da história da escrita e da leitura, e da história das mulheres

no Brasil, por meio da construção de um catálogo de documentos escritos por mulheres e documentos escritos sobre mulheres (contendo sua ‘fala’ na forma de discurso relatado) na América Portuguesa. O catálogo e o tratamento dos textos no M.A.P. são desenvolvidos como recursos digitais que buscam unir tecnologias autais aos mais rigorosos princípios da prática filológica, nos moldes de Paixão de Sousa, 2013 e Monte e Paixão de Sousa, 2017.

O LaPeLinC (<http://memoriaconquistense.uesb.br/websinc/>) tem sua história iniciada também em 2009, na Universidade Estadual do Sudoeste da Bahia, campus de Vitória da Conquista, com a elaboração do projeto Memória Conquistense, pelos Professores Dr. Jorge Viana Santos e Dr^a Cristiane Namiuti-Temponi. Atualmente, o laboratório desenvolve, entre outros, os projetos Memória conquistense: implementação de um corpus digital, coordenado pela pela Prof^a. Dr^a Cristiane Namiuti Temponi (Namiuti, 2013), e Corpora digitais de documentos históricos da imperial Vila da Victoria, atual Vitória da Conquista - Bahia: resgate e preservação do patrimônio linguístico e da memória da escravidão na Bahia, coordenado pelo Prof. Dr. Jorge Viana Santos (Santos, 2016). Os Projetos desenvolvidos no LaPeLinC têm focalizado a construção de corpora digitais que partem de documentos históricos, manuscritos e impressos, de diferentes gêneros, guardados em arquivos públicos e/ou particulares da Bahia, da região Sudoeste do estado e da cidade de Vitória da Conquista, visando a disponibilização de textos antigos no formato singular do “texto digital”, guiados pelo critério básico da transparência da codificação (cf. Santos e Namiuti, 2017, 2018, 2019).

Nesses dez anos desde a criação do LaPeLinC e do Grupo de Pesquisas Humanidades Digitais, seus pesquisadores participaram em conjunto em diversas atividades ligadas ao desenvolvimento de corpora e à anotação linguística. As Prof^{as}. Maria Clara Paixão de Sousa e Cristiane Namiuti vêm participando conjuntamente em diversos encontros da área, tais como o Workshop Construction and use of large annotated corpora (Unicamp, 2013, <http://www.tycho.iel.unicamp.br/ws2013>); o II Encontro de Gramática Gerativa – Relações multidisciplinares e mudança linguística, promovido pela UESB em 2018 (<https://iiencontrogramaticagerativa.wordpress.com>), e a III Jornada Científica do Lapelinc/I Workshop em Anotação e Frameworks para Corpora Digitais de Textos Históricos, em 2019 (<https://iiijornadalapelinc.wordpress.com>), entre outros. Além disso, desde 2009, a Prof^a. Maria Clara é colaboradora nos projetos de pesquisa do LaPeLinC (atualmente, Namiuti, 2016; Santos, 2016). Por fim, ressalte-se que, entre 1998 e 2007, as coordenadoras fizeram parte da equipe fundadora do Corpus Tycho Brahe (o primeiro corpus sintaticamente anotado do Português, cf. <http://www.tycho.iel.unicamp.br/corpus/>), um grupo de pesquisas ativo por mais de vinte anos, pioneiro no uso de ferramentas computacionais em linguística e precursor das Humanidades Digitais no Brasil (cf. Namiuti, 2004; Paixão de Sousa, 2014).

6 Metodologia

O ambiente virtual do LaViHD seguirá os moldes do WebSinc (Namiuti, Santos e Costa, 2015; Costa, 2015) do LaPeLinC, um software desenvolvido para o trabalho de disponibilização, visão e busca de dados em corpora cientificamente controlados e anotados em diversos níveis. O WebSinc inclui módulos de Sistemas de Gerenciamento de Informação (SGDI), bancos de dados, e ferramentas de buscas (morfossintáticas e sintáticas) em padrão XML, como ilustram a Figura 1 abaixo e as Figuras 4 a 23 no Anexo.



Figura 1 - Exemplo de uma instância do Catálogo Visual do WebSinc, com a representação parte imagética e descrição básica de um livro manuscrito catalogado (Livro de Notas 14, 1882-1883)



Figura 2: Visão parcial da visualização georreferenciada do Catálogo M.A.P.

A essas ferramentas desenvolvidas no WebSinc serão acrescentadas funcionalidades em desenvolvimento no Projeto M.A.P. para maior flexibilidade na visualização dos dados – em particular, com o georreferenciamento dos documentos catalogados (cf. Figura 2). É relevante ressaltar que também neste Projeto a linguagem base é o XML, o que demonstra a convergência entre as metodologias dos grupos (e, portanto, dá efetividade ao intercâmbio planejado).

Por fim, figura como elo fundamental entre os objetivos tecnológicos dos dois projetos o desenvolvimento do eDictor 2.0, a partir do eDictor 1.0 (cf. Figura 3), um software para edições filológicas eletrônicas criado pelo Grupo Humanidades Digitais e atualmente em uso pelas duas equipes – além de diversas outras, no Brasil e em Portugal (cf. <https://edictor.net>, e Paixão de Sousa 2013, 2019).

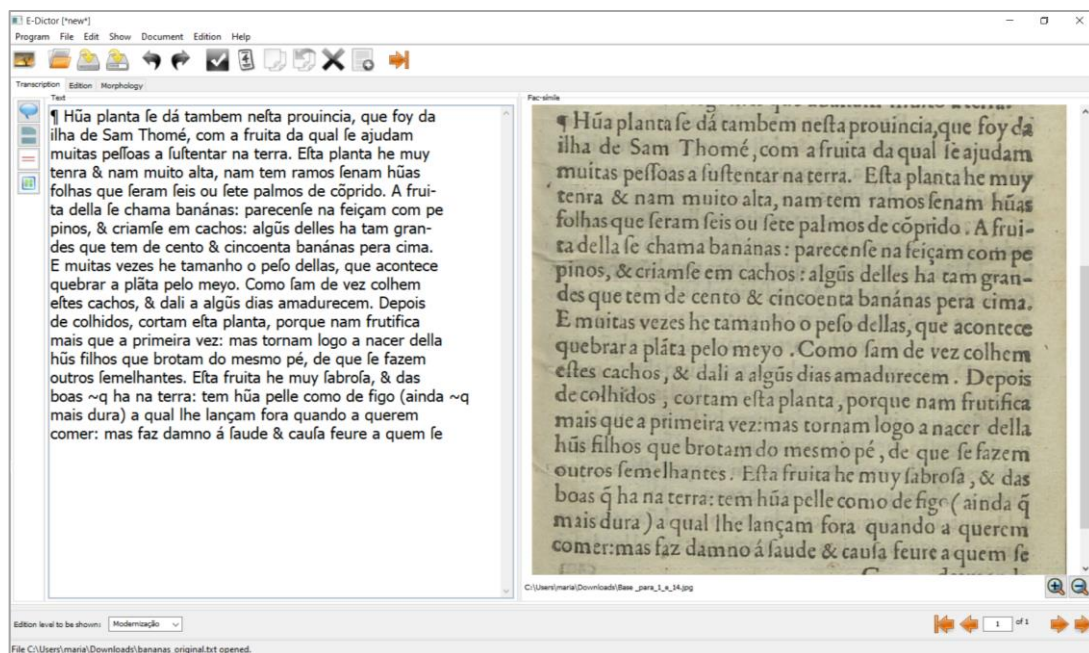


Figura 3 - eDictor 1.0 - Módulo de transcrição filológica

O eDictor 2.0 será inteiramente baseado na Web, e seu desenvolvimento terá por objetivo, a um tempo, sanar algumas deficiências conhecidas da versão 1.0 (Paixão de Sousa, 2015) e intensificar seus pontos fortes. Esse desenvolvimento está já em fase preliminar, por iniciativa de uma das pesquisadoras do LaPeLinC, Aline Silva Costa, doutoranda do Programa de Pós-graduação em Linguística da UESB, sob orientação da Prof^a. Cristiane Namiuti e coorientação da Prof^a. Maria Clara Paixão de Sousa (Costa, 2019, em curso).

7 Cronograma

O objetivo imediato do Projeto é implementar o ambiente virtual de laboratório para uso conjunto dos pesquisadores das duas instituições parceiras (LaViHD^{beta}) e, no longo prazo, abrir uma versão estável do ambiente para amplo uso pela comunidade de pesquisas. A meta é terminar o desenvolvimento do ambiente fechado até o final de **2022**, e lançar o ambiente estável e aberto em **2023**.

	2020	2022	2023
Desenvolvimento do LaViHD ^{beta} e testes em uso interno	✓	✓	
Desenvolvimento do LaViHD estável		✓	
Lançamento do LaViHD estável (aberto)			✓

8 Referências bibliográficas

- Costa, AS. Um sistema de anotação de múltiplas camadas para o corpus digital DOViC [tese]. Tese de doutorado, Departamento de Estudos Linguísticos e Literários, Universidade Estadual do Sudoeste da Bahia. Vitória da Conquista; em curso.
- Costa, AS. WebSinC: Uma Ferramenta Web para buscas sintáticas e morfossintáticas em corpora anotados - Estudo de Caso do Corpus DOViC – Bahia. Dissertação de Mestrado, Departamento de Estudos Linguísticos e Literários, Universidade Estadual do Sudoeste da Bahia. Vitória da Conquista; 2015.
- Monte, VM do. M.A.P. (Mulheres na América Portuguesa): Mapeamento digital de escritos de mulheres e sobre mulheres no espaço atlântico português – Vertente filológica. Projeto de pesquisa, Programa Unificado de Bolsas (PUB) / Universidade de São Paulo. São Paulo; 2018.
- Monte, VM do; Paixão de Sousa, MC. Agora andam me jurando a pele: Escritos de mulheres e escritos sobre mulheres na América Portuguesa. Projeto de pesquisa, Programa Unificado de Bolsas (PUB), Universidade de São Paulo. São Paulo; 2017(a). URL: http://www.nehlp.org/~nehlp/HD/MAP/MAP_Projeto_2017.html
- Monte, VM do; Paixão de Sousa, MC. Por uma filologia virtual: O caso das atas da câmara de São Paulo (1562-1596). REVISTA DA ABRALIN, v. 16, p. 239-264, 2017(b). URL: <http://revistas.ufpr.br/abralin/article/view/51938>
- Namiuti, C, Santos, JV; Costa, Aline Silva. 2015. WebSinC. Vitória da Conquista: Universidade Estadual do Sudoeste da Bahia; 2015. URL: <http://memoriaconquistense.uesb.br/websinc>
- Namiuti, C. Corpora digitais para a história do Português Brasileiro - Documentos históricos da região sudoeste da Bahia: Aliança PHPP-Tycho Brahe. Projeto de pesquisa. Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB 6171/2010) / Universidade Estadual do Sudoeste da Bahia. Vitória da Conquista; 2010.
- Namiuti, C. Do português pré-clássico às variantes modernas: contribuições para o estudo da sintaxe e suas interfaces. Projeto de pesquisa. Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB 007/2016) / Universidade Estadual do Sudoeste da Bahia. Vitória da Conquista; 2016.
- Namiuti, C. Memória conquistense: implementação de um corpus digital. Projeto de pesquisa. Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, 485098/2013-0) / Universidade Estadual do Sudoeste da Bahia. Vitória da Conquista; 2013.
- Namiuti, C. O corpus anotado do português histórico: um avanço para as pesquisas em lingüística histórica do português. Revista Virtual de Estudos da Linguagem, www.revelhp.cjb.net, v. 3, p. 1-9, 2004.
- Paixão de Sousa, MC. A Filologia Digital em Língua Portuguesa: Alguns caminhos. In: Ana Paula Banza & Maria Filomena Gonçalves (coord.), Património textual e humanidades digitais: da antiga à nova Filologia. Évora: CIDEHUS/FCT; 2013. URL: <http://dspace.uevora.pt/rdpc/bitstream/10174/10468/1/e-book.pdf>
- Paixão de Sousa, MC. eDictor: A chronology. Apresentação na mesa redonda “eDictor: advances and perspectives”. Workshop Construction and use of large annotated corpora. Campinas; 9 de setembro de 2013. URL: <http://manualedictor.wordpress.com/historico>

Paixão de Sousa, MC. e-Dictor: Histórico e perspectivas. Comunicação ao Gallæcia: III Congresso Internacional de Linguística Histórica. Santiago de Compostela; 28 de julho 2015. URL: <http://www.slideshare.net/mariaclaraps/edictor-historico-e-perspectivas-2015>

Paixão de Sousa, MC. Memórias do texto (2004-2019): trajetórias, desvios e novos caminhos. Conferência na III Jornada Científica do Lapelinc / I Workshop em Anotação e Frameworks para Corpora Digitais de Textos Históricos; 2019.

Paixão de Sousa, MC. M.A.P. (Mulheres na América Portuguesa): Mapeamento digital de escritos de mulheres e sobre mulheres no espaço atlântico português. Projeto de pesquisa, Programa Unificado de Bolsas (PUB) / Universidade de São Paulo. São Paulo; 2018.

Paixão de Sousa, MC. O Corpus Tycho Brahe: contribuições para as humanidades digitais no Brasil. Filologia e Linguística Portuguesa, v. 16, p. 53-93, 2014. URL: <http://www.revistas.usp.br/flp/article/view/88404>

Paixão de Sousa, MC; Kepler, FN. E-Dictor: Uma ferramenta integrada para a anotação de edição e classe de palavras. VI Encontro de Linguística de Corpus, São Paulo; 2007.

Paixão de Sousa, MC; Kepler, FN; Faria, PPF de. E-Dictor: Novas perspectivas na codificação e edição de corpora de textos históricos. In: Shepherd T; Sardinha T; Pinto MV., organizadores. Caminhos da linguística de corpus. Campinas: Mercado de Letras; 2010.

Paixão de Sousa, MC; Kepler, FN; Faria, PPF de. e-Dictor. Versão 1.0 beta 10, 2013. Programa de Computador. URL: <https://edictor.net/download>

Paixão de Sousa, MC; Monte, VM do. M.A.P. (Mulheres na América Portuguesa): Mapeamento de escritos de mulheres e sobre mulheres no espaço atlântico português a partir de métodos das Humanidades Digitais. Projeto de pesquisa, Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq, Edital Universal 2018) / Universidade de São Paulo. São Paulo; 2018.

Santos, JV. Corpora digitais de documentos históricos da imperial Vila da Victoria, atual Vitória da Conquista - Bahia: resgate e preservação do patrimônio linguístico e da memória da escravidão na Bahia. Projeto de pesquisa. Fundação de Amparo à Pesquisa do Estado da Bahia (FAPESB 014/2016) / Universidade Estadual do Sudoeste da Bahia. Vitória da Conquista; 2016.

Santos, JV; Namiuti, C. De manuscritos históricos a corpora anotados: do Documento Físico (DF) ao Documento Digital Imagem (DDI). In: Barreiros, PN; Almeida, IS; Santos, RB, organizadores. Filologia e Humanidades Digitais. Feira de Santana: Editora UEFS; 2018.

Santos, JV; Namiuti-Temponi, C. Memória Conquistense. UESB/LAPELINC, Vitória da Conquista; 2016. URL: <http://memoriaconquistense.uesb.br/websinc>

Santos, JV; Namiuti-Temponi, C. O futuro das Humanidades Digitais é o passado. In: Carrilho, E, Martins, AM, Pereira S; Silvestre JP, organizadores. Estudos Linguísticos e Filológicos Oferecidos a Ivo Castro. Lisboa: Centro de Linguística da Universidade de Lisboa; 2019.

Santos, JV; Namiuti-Temponi, C. O objeto livro: a complexidade da forma e o digital Anais do X Congresso Internacional da ABRALIN; Universidade Federal Fluminense; 2017.

Anexo

Ilustrações da interface do WebSinc para catalogação e consultas

A. Catalogação: interface para cadastro de documentos

Figura 4 - Cadastro da tipologia dos Documentos Macro no WebSinc

Figura 5 - Cadastro de Documento Macro

Figura 6 - Cadastro de Documento Macro (ii) – Tipologia em uso no WebSinc

Figura 7 - Cadastro de Documento Macro – Tela de inserção dos dados

Figura 8 - Cadastro de Folhas-imagem do Documento Macro cadastrado

Figura 9 - Cadastro de Folhas-imagem do Documento Macro cadastrado: Tela de inserção da Folha-imagem

Figura 10 - Cadastro de Manuscritos (contidos e topografados em cada Documento Macro): Dados Gerais

Figura 11 - Cadastro de Manuscritos (contidos e topografados em cada Documento Macro): Arquivos Vinculados

Figura 12 - Cadastro de Manuscritos (contidos e topografados em cada Documento Macro): Folhas-imagem vinculadas

Figura 13 - Cadastro de Manuscritos (contidos e topografados em cada Documento Macro): Lista final de documentos cadastrados em um mesmo livro (Documento Macro)

B. Interface para consultas e buscas

Figura 14 - Consulta ao Catálogo – Lista dos Documentos Macro para consulta

Figura 15 - Consulta ao Catálogo – Exemplo de um Documento Macro com os dados imagéticos e tipológicos, e (abaixo) a lista dos Documentos Micro nele contidos.

Figura 16 - Consulta ao Catálogo – Documento Micro com a transcrição filológica conservadora

Figura 17 - Consulta ao Catálogo – Documento Micro com a transcrição filológica modernizada

Figura 18 - Consulta ao Catálogo – Glossário de edições em um Documento Micro

Figura 19 - Consulta ao Catálogo – Tela de seleção de um documento para busca linguística

Figura 20 - Consulta ao Catálogo – Resultado de busca morfossintática em um manuscrito

Figura 21 - Consulta ao Catálogo – Resultado de busca sintática (lista de ocorrências)

Figura 22 - Consulta ao Catálogo – Resultado de busca sintática (árvore sintagmática horizontal)

Figura 23 - Consulta ao Catálogo – Resultado de busca sintática (árvore sintagmática vertical)

A. Catalogação: interface para cadastro de documentos


Início Cadastros Corpus Catálogo Buscas Relatórios Configurações Sobre Sair	
Tipos de Documento Macro	
+ Novo	
Descrição do tipo	Ações
Contratos	 
Escritura de cessão de direitos hereditários	 
Escritura de convenção com pacto antenupcial	 
Escritura de reconhecimento de filhos ilegítimos	 
Livro de busca	 
Livro de escrituras	 
Livro de notas	 
Livro de procurações	 
Obra Literária	 
Processos	 

Figura 4 - Cadastro da tipologia dos Documentos Macro no WebSinc

Início

Cadastros

Corpus

Catálogo

Buscas

Relatórios

Configurações

Sobre

Sair

Cadastro de Documento Macro

+ Novo

Início

























Código	Título do documento	Ano	Tipo de documento	Vinculados ao documento	Ações
CTB	Corpus Anotado do Português Histórico Tycho Brahe - PCI	1679	Obra Literária	ImagensManuscritos	 
EA1-1	Escritura de convenção com pacto antenupcial nº1	1982	Escritura de convenção com pacto antenupcial	ImagensManuscritos	 
LB1-1	Livro de Busca nº1	1919	Livro de busca	ImagensManuscritos	 
LB7-10	Livro de Busca nº10	1977	Livro de busca	ImagensManuscritos	 
LB6-10	Livro de Busca nº10	1977	Livro de busca	ImagensManuscritos	 
LB8-11	Livro de Busca nº11	1980	Livro de busca	ImagensManuscritos	 
LB9-12	Livro de Busca nº12		Livro de busca	ImagensManuscritos	 
LB10-13	Livro de Busca nº13	1985	Livro de busca	ImagensManuscritos	 
LB11-14	Livro de Busca nº14		Livro de busca	ImagensManuscritos	 
LB12-15	Livro de Busca nº15		Livro de busca	ImagensManuscritos	 
LB13-16	Livro de Busca nº16		Livro de busca	ImagensManuscritos	 
LB2-2	Livro de Busca nº2	1935	Livro de busca	ImagensManuscritos	 
LB3-3	Livro de Busca nº3	1949	Livro de busca	ImagensManuscritos	

Figura 5 - Cadastro de Documento Macro

Início Cadastros Corpus Catálogo Buscas Relatórios Configurações Sobre Sair					
Cadastro de Documento Macro					
+ Novo Início					
Código	Título do documento	Ano	Tipo de documento	Vinculados ao documento	Ações
CTB	Corpus Anotado do Português Histórico Tycho Brahe - PCI	1679	Todos os tipos		
EA1-1	Escritura de convenção com pacto antenupcial nº1	1982	Contratos		
LB1-1	Livro de Busca nº1	1919	Escritura de cessão de direitos hereditários		
LB7-10	Livro de Busca nº10	1977	Escritura de convenção com pacto antenupcial		
LB6-10	Livro de Busca nº10	1977	Escritura de reconhecimento de filhos ilegítimos		
LB8-11	Livro de Busca nº11	1980	Livro de busca		
LB9-12	Livro de Busca nº12		Livro de escrituras		
LB10-13	Livro de Busca nº13	1985	Livro de notas	Imagens Manuscritos	
LB11-14	Livro de Busca nº14		Livro de procurações	Imagens Manuscritos	
LB12-15	Livro de Busca nº15		Obra Literária	Imagens Manuscritos	
LB13-16	Livro de Busca nº16			Imagens Manuscritos	
LB2-2	Livro de Busca nº2	1935		Imagens Manuscritos	
LB3-3	Livro de Busca nº3	1949		Imagens Manuscritos	

Figura 6 - Cadastro de Documento Macro (ii) – Tipologia em uso no WebSinc

Início Cadastros Corpus Catálogo Buscas Relatórios Configurações Sobre Sair											
Cadastro de Documento Macro											
+ Novo Cancelar											
Código:	*Título:	* Tipo de Documento:		Sel. tipo							
Descrição do Documento:											
Local de depósito:											
Arquivo do 1º. Tabelionato de notas Paes (antigo 1º Cartó)											
Tipo de data:	Data de início	Data de fim	Ano inicial	Ano final							
Selecione:											
Altura (cm):	Largura (cm):	Profundidade (cm):									
Com capa	Capa original	Com forro									
Material da capa:		Cor da capa:									
<input type="radio"/> Couro <input type="radio"/> Papelão <input type="radio"/> Plástico <input type="radio"/> Tecido		<input type="radio"/> Couro <input type="radio"/> Papelão <input type="radio"/> Plástico <input type="radio"/> Tecido									
Cor da capa:		Sel. cor:									
Observações:											
Cancelar Salvar											

Figura 7 - Cadastro de Documento Macro – Tela de inserção dos dados

Início Cadastros Corpus Catálogo Buscas Relatórios Configurações Sobre Sair							
Cadastro de Imagens do Documento Macro							
Documento Macro: Livro de Notas 14 dos annos de 1882 a 1883							
+ Novo Voltar							
Imagem	Nº	Código	Tipo	Lado	Nº Folha	Nº Página	Ações
	0	E14-C0-null	LOMBADA	FRENTE	0	0	
	1	E14-C11-1	CAPA	FRENTE	0	0	
	2	E14-C11-2	CONTRA CAPA	VERSO	0	0	
	3	E14-C11-3	TERMO DE ABERTURA	FRENTE	0	0	
	4	E14-C11-4	FOLHA	VERSO	0	0	
	5	E14-C11-5	FOLHA	FRENTE	1	0	
	6	E14-C11-6	FOLHA	VERSO	1	0	
	7	E14-C11-7	FOLHA	FRENTE	2	0	
	8	E14-C11-8	FOLHA	VERSO	2	0	
	9	E14-C11-9	FOLHA	FRENTE	3	0	

Figura 8 - Cadastro de Folhas-imagem do Documento Macro cadastrado

Cadastro de Imagens do Documento Macro			
Documento Macro: Livro de Notas 14 dos annos de 1882 a 1883			
+ Novo			
Imagem	Nº	Código	Tipo
	0	E14-C0-null	LOMBADA
	1	E14-C11-1	CAPA
	2	E14-C11-2	CONTRA CAPA
	3	E14-C11-3	TERMO DE ABERTURA
	4	E14-C11-4	FOLHA
	5	E14-C11-5	FOLHA
	6	E14-C11-6	FOLHA
	7	E14-C11-7	FOLHA
	8	E14-C11-8	FOLHA
	9	E14-C11-9	FOLHA

Imagens

Código da imagem: E14-C0-1

Nº da coleta: Nº de sequência: Nº da página: Nº da folha:

1 0 0

Tipo de Imagem:

☐ FOLHA

☒ CAPA

☐ CAPA INTERNA

☐ CONTRA CAPA

☐ LOMBADA

☐ TERMO DE ABERTURA

☐ TERMO DE FECHAMENTO

Lado da Imagem:

☒ FRENTE

☐ VERSO

Selecionar arquivo do computador:

+ Selecionar

IMG_1.494334720018E12.jpg

Cancelar Salvar

Figura 9 - Cadastro de Folhas-imagem do Documento Macro cadastrado: Tela de inserção da Folha-imagem

Inserção de Manuscritos

Manuscritos do Documento Livro de Notas 1 dos annos de 1841 a 1848

+ Novo [Cancelar](#)

Dados Gerais
Arquivos Vinculados
Imagens do Documento
Créditos - Transcrição
Créditos - Edição
Créditos - Morfologia
Créditos - Sintaxe
Créditos-Cadastro

Dados Gerais do Documento
Código: 1.37 *Título: Carta de Liberdade de Adrianna Mulata * Tipo de Documento: Carta de alforria
Data do documento 31/01/1845 Data de publicação Ano do documento Ano de publicação
Observações
* Campos obrigatórios
[Cancelar](#) [Salvar](#)

Figura 10 - Cadastro de Manuscritos (contidos e topografados em cada Documento Macro): Dados Gerais

Inserção de Manuscritos

Manuscritos do Documento Livro de Notas 1 dos annos de 1841 a 1848

+ Novo [Cancelar](#)

Dados Gerais
Arquivos Vinculados
Imagens do Documento
Créditos - Transcrição
Créditos - Edição
Créditos - Morfologia
Créditos - Sintaxe
Créditos-Cadastro

Arquivos anotados
Arquivo XML - Anotação POS
Válido apenas arquivo com extensão .xml):
+ Selecionar
Arquivo gravado no servidor:
Arquivo XML: XML_1.508952097612E12.xml [Download](#) Excluir
Arquivo PSD - Anotação sintática
Válido apenas arquivo com extensão .txt):
+ Selecionar
*Um arquivo sintático no formato XML será automaticamente gerado a partir do arquivo PSD
Arquivos gravados no servidor:
Arquivo PSD: [Download](#) Excluir
Arquivo XML: [Download](#) Excluir
* Campos obrigatórios
[Cancelar](#) [Salvar](#)

Figura 11 - Cadastro de Manuscritos (contidos e topografados em cada Documento Macro): Arquivos Vinculados

[Início](#)
[Cadastros](#)
[Corpus](#)
[Catálogo](#)
[Buscas](#)
[Relatórios](#)
[Configurações](#)
[Sobre](#)
[Sair](#)

Inserção de Manuscritos

Manuscritos do Documento Livro de Notas 1 dos anos de 1841 a 1848

[+ Novo](#)
[Cancelar](#)

Dados Gerais

Arquivos Vinculados

Imagens do Documento

Créditos - Transcrição

Créditos - Edição

Créditos - Morfologia

Créditos - Sintaxe

Créditos-Cadastro

Imagens do Documento

Imagem	Nº	Código	Tipo	Numeração	Ações
	208	E1-C11-208	FOLHA	Número Folha:101 Número Página:0	
	209	E1-C11-209	FOLHA	Número Folha:102 Número Página:0	
	210	E1-C11-210	FOLHA	Número Folha:102 Número Página:0	

[+ Adicionar Imagens](#)

* Campos obrigatórios

[Cancelar](#)
[Salvar](#)

Figura 12 - Cadastro de Manuscritos (contidos e topografados em cada Documento Macro): Folhas-imagem vinculadas

[Início](#)
[Cadastros](#)
[Corpus](#)
[Catálogo](#)
[Buscas](#)
[Relatórios](#)
[Configurações](#)
[Sobre](#)
[Sair](#)

Inserção de Manuscritos

Manuscritos do Documento Livro de Notas 14 dos anos de 1882 a 1883

[+ Novo](#)
[Voltar](#)

Código	Título	Ano	Tipo de documento	Ações
45	Carta de Liberdade do escravo Francisco	1883	Carta de alforria	
7	Procuração bastante que faz Manoel Pedro do Carmo como abaixo se declara.		Procuração	
4	Procuração bastante que faz Anselmo da Silva Brasileiro a seos procuradores abaixo declarado.		Procuração	
5	Procuração bastante que faz Anselmo da Silva Brasileiro a seos procuradores abaixo declarados.		Procuração	
6	Procuração bastante que faz Dona Maria Rosa da Ponte a seos procuradores abaixo declarados.		Procuração	
8	Procuração bastante que faz Dona Virginia Emerentina da Silveira, professora pública desta Imperial Vila da Victoria a seu procurador abaixo declarado.		Procuração	
9	Procuração bastante que fazem Felisberto Pereira de Oliveira, Alfredo Honorato Xavier Soares, Alfredo Xavier Soares a seus procuradores abaixo declarados.		Procuração	
10	Procuração bastante que faz Francisco Febroseio? Alves dos Reis a seos procuradores abaixo declarados.		Procuração	
54	Carta de liberdade [?]		Carta de alforria	
23	Procuração bastante que faz o Professor Publico Antonio Francisco de Carvalho a seu procurador abaixo declarado.		Procuração	

Figura 13 - Cadastro de Manuscritos (contidos e topografados em cada Documento Macro): Lista final de documentos cadastrados em um mesmo livro (Documento Macro)

B. Interface para consultas e buscas

Início	Cadastros	Corpus	Catálogo	Buscas	Relatórios	Configurações	Sobre	Sair
------------------------	---------------------------	------------------------	--------------------------	------------------------	----------------------------	-------------------------------	-----------------------	----------------------

Catálogo Corpus DeVic			
Documentos macro (livros):			
Id	Título	Ano	Tipo de documento
			Livro de notas
E14	Livro de Notas 14 dos annos de 1882 a 1883	1882	Livro de notas
E16	Livro de Notas 16 dos annos de 1883 a 1884	1883	Livro de notas
E1	Livro de Notas 1 dos annos de 1841 a 1848	1841	Livro de notas
E5	Livro de Notas 5 dos annos de 1862 a 1866	1862	Livro de notas
E2	Livro de Notas dos annos de 1841 a 1855	1841	Livro de notas
E3	Livro de Notas dos annos de 1849 a 1858	1849	Livro de notas
E6	Livro de Notas dos annos de 1869 a 1870	1869	Livro de notas
E8	Livro de Notas dos annos de 1870 a 1874	1870	Livro de notas
E11	Livro de Notas dos annos de 1877 a 1880	1877	Livro de notas
E13	Livro de Notas dos annos de 1880 a 1881	1880	Livro de notas
E21	Livro de Notas dos annos de 1880 a 1890	1880	Livro de notas
E15	Livro de Notas dos annos de 1881 a 1882	1881	Livro de notas
E22	Livro de Notas dos Annos de 1891 a 1900	1891	Livro de notas
E7	Livros de Notas dos annos de 1864 a 1871	1864	Livro de notas

Figura 14 - Consulta ao Catálogo – Lista dos Documentos Macro para consulta



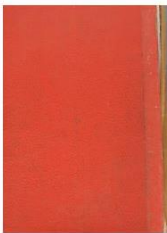


INFORMAÇÕES Livro de Notas 1 dos annos de 1841 a 1848				
Título: Livro de Notas 1 dos annos de 1841 a 1848 Tipo: Livro de notas Ano: 1841 - 1848 Capa: Papelão - Marrom Altura: 31,5 cm Largura: 22,3 cm Profundidade: 3,0 cm				
CAPA (FRENTE)	LOMBADA	CAPA (VERSO)	TERMO DE ABERTURA	TERMO DE FECHAMENTO
				
Documentos do Livro de Notas 1 dos annos de 1841 a 1848				
Título				Ano do document
Carta de liberdade da Cabra de nome Sofia				1841
Carta de Liberdade de Manoel Africano				1845
Carta de Liberdade do Cabrinha Remardo				1845
Carta de Liberdade de Adrianna Mulata				

Figura 15 - Consulta ao Catálogo – Exemplo de um Documento Macro com os dados imagéticos e tipológicos, e (abaixo) a lista dos Documentos Micro nele contidos.



Figura 16 - Consulta ao Catálogo – Documento Micro com a transcrição filológica conservadora



Figura 17 - Consulta ao Catálogo –Documento Micro com a transcrição filológica modernizada



Figura 18 - Consulta ao Catálogo – Glossário de edições em um Documento Micro



Figura 19 - Consulta ao Catálogo – Tela de seleção de um documento para busca linguística

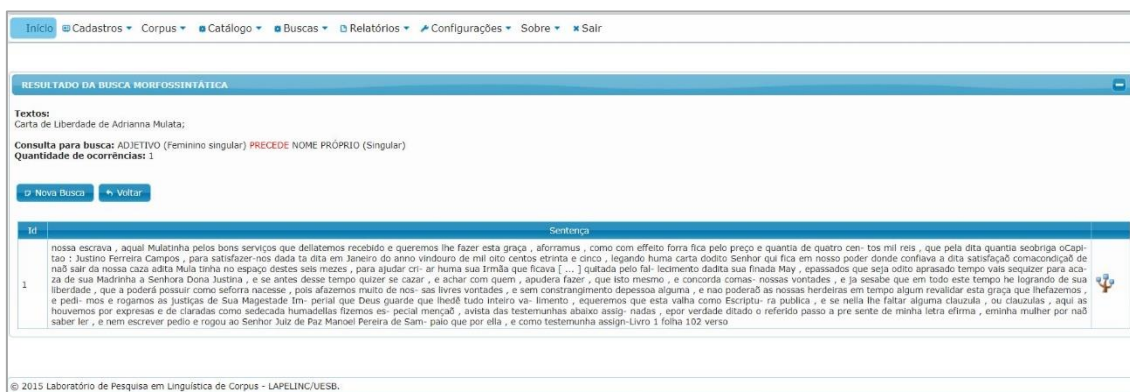


Figura 20 - Consulta ao Catálogo – Resultado de busca morfossintática em um manuscrito



Figura 21 - Consulta ao Catálogo – Resultado de busca sintática (lista de ocorrências)

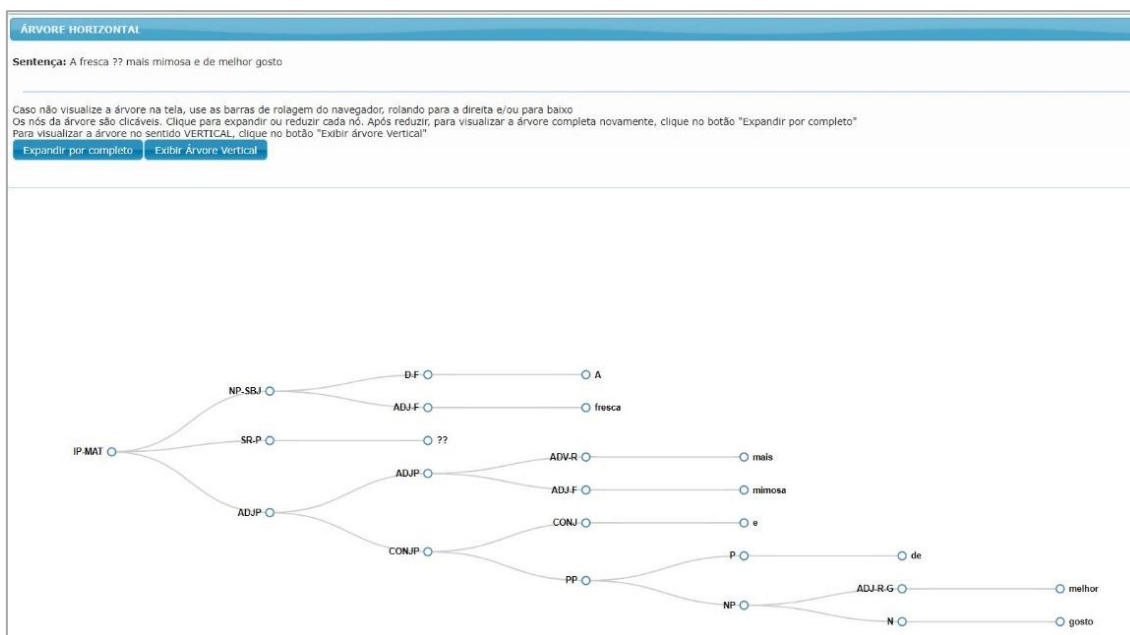


Figura 22 - Consulta ao Catálogo – Resultado de busca sintática (árvore sintagmática horizontal)

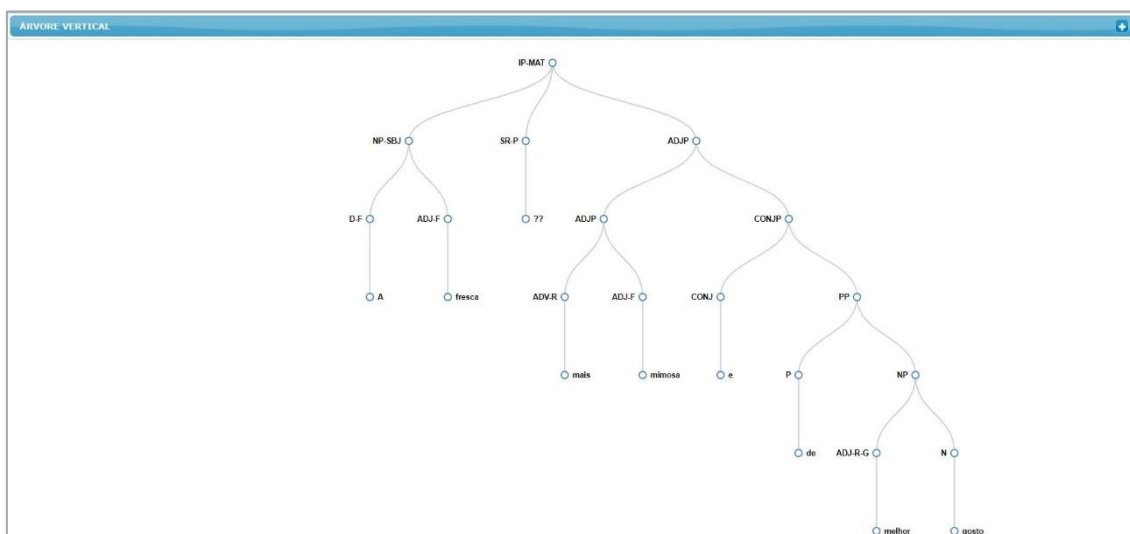


Figura 23 - Consulta ao Catálogo – Resultado de busca sintática (árvore sintagmática vertical)