Corpus M.A.P.:

Manual de Edição



Manual de edição filológica

Projeto M.A.P.: Mulheres na América Portuguesa

https://map.prp.usp.br

Coordenação

Maria Clara Paixão de Sousa Vanessa Martins do Monte

Equipe

Andrea Cristina Natanael da Silva Elisa Hardt Leitão Motta Nicólli de Lima Garcia Priscila Starline Estrela Tuy Batista

O Corpus M.A.P.: Desenvolvimento de um sistema de edições filológicas digitais para a documentação sobre Mulheres na América Portuguesa. Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, Auxílio à Pesquisa 2021/15133-2.

Como referenciar.

Paixão de Sousa, Maria Clara; Monte, Vanessa Martins do; Garcia, Nicólli de Lima; Motta, Elisa Hardt Leitão; Silva, Andrea Cristina Natanael da; Tuy Batista, Priscila Starline Estrela. **Manual de descrição documental: Projeto Mulheres na América Portuguesa**. Universidade de São Paulo, 2025. https://map.prp.usp.br/Manual/MAP_descricao.html

Sumário

Apresentação	4
1. Processo de edição de texto 1.0 Funcionamento básico do eDictor	5 5
1.1 Edição de texto e etiquetas	7
1.1.1 Éfeitos estruturais e resumo	11
1.2 Descrição de elementos de texto	12
1.2.1. Seção	16
1.2.2. Parágrafo	17
1.2.3. Sentença	18
1.2.4 Tipo de palavra	18
1.2.5 Como fazer anotações	25
1.2.6 Efeitos estruturais e resumo	28
1.3 Metadados	28
1.3.1 Metadados originais do eDictor	28
1.3.2 Uso customizado dos metadados do eDictor no Co M.A.P.	rpus 30
1.4 Apresentação final das edições	33
1.4.1 Visualizações das camadas de edição	33
1.4.2 Aparência e estilos	34
2. Normas de edição	35
2.0 Etapa de Transcrição	35
2.1. Camada diplomática	37
2.2 Camada semidiplomática	39
2.3 Camada atualizada	41
2.4 Camada modernizada	43
2.5 Versões para visualização	46
4. Como referenciar	46
4.1 O Corpus M.A.P	46
4.2 As edições eletrônicas	46
eDictor: tutorial de instalação, customização e uso básico	47
1. Download e instalação	48
2. Módulo de transcrição	51
3 Módulo de edição	53

Apresentação

Este manual mostra as normas de edição filológica seguidas no Corpus M.A.P. - Mulheres na América Portuguesa e desenvolvidas pela equipe do Projeto O Corpus M.A.P.: desenvolvimento de um sistema de edições filológicas digitais para a documentação sobre Mulheres na América Portuguesa¹.

O trabalho de edição no Corpus é realizado com o uso do software **eDictor** (https://edictor.net), um editor de textos desenvolvido para o trabalho filológico e a análise linguística automática. A metodologia seguida pela equipe do Projeto procura combinar técnicas e princípios da edição filológica tradicional com as possibilidades e requerimentos técnicos deste software. Em particular, isso significa que os resultados do trabalho de edição ficam reunidos em um só documento digital que pode gerar diferentes versões para leitura.

Tradicionalmente, os termos pelos quais referenciamos os "níveis de edição" filológica -"edição diplomática", "edição semidiplomática" e "edição modernizada" - são usados para designar tanto os diferentes *processos* de edição como os diferentes *produtos* da edição. Em nosso sistema, é preciso diferenciar esses dois sentidos: ao realizar a edição, estamos aplicando sempre todos os níveis de intervenção; ao dar a edição à leitura, podemos selecionar diferentes níveis de intervenção. Assim, a rigor, a "edição diplomática", a "edição semidiplomática" e a "edição modernizada" de um documento trabalhado neste método são diferentes **versões de leitura** de **uma hiper-edição** em camadas.

Assim, a principal diferença substantiva entre o método do eDictor e o trabalho de edição tradicional é que a relação entre a 'edição' enquanto **processo** e a 'edição' enquanto **produto** é reorganizada: *o processo de edição* de um documento envolve diferentes graus de intervenção, todos reunidos em um só arquivo, de onde podem ser gerados diferentes *produtos de edição* no sentido de diferentes possibilidades de visualização das intervenções editoriais. Chamamos a isso uma **edição em camadas**.

Neste Manual, procuramos mostrar esse método em duas partes: primeiro, mostramos os procedimentos que seguimos ao longo do trabalho de edição - da transcrição do texto à inserção e codificação das intervenções editoriais - e, em seguida, mostramos os diferentes produtos que podem ser gerados a partir desses procedimentos realizados em camadas.

Por isso, usaremos os termos "camada diplomática", "camada semidiplomática", "camada atualizada", "camada modernizada" para designar cada camada de intervenção realizada no *processo* de produção da edição. Por conformidade com a nomenclatura tradicional, usaremos os termos "edição diplomática", "edição semidiplomática" e "edição modernizada" no sentido de *produtos* do trabalho editorial.

¹ Financiado pela Fundação de Amparo à Pesquisa do Estado de São Paulo – FAPESP, na modalidade Auxílio à Pesquisa, sob o número FAPESP 2021/15133-2

1. Processo de edição de texto

1.0 Funcionamento básico do eDictor

O eDictor é um editor de textos desenvolvido para o trabalho filológico e a análise linguística automática. Ao combinar um editor de XML e um etiquetador morfossintático, a ferramenta permite que sejam geradas, automaticamente, diferentes versões de edição para leitura: diplomáticas, semidiplomática e modernizadas, além de versões com anotação morfossintática (em XML e texto simples) (Paixão de Sousa, 2013). O eDictor é o programa adotado pelo Projeto M.A.P. e outros seis projetos de pesquisa². O software pode ser baixado em (https://edictor.net); tutorial para sua instalação pode ser encontrado em о https://manualeDictor.wordpress.com/eDictor/, e um roteiro para seu uso básico е personalização no M.A.P. está disponível no final deste manual, no Apêndice eDictor: tutorial de instalação, customização e uso básico.

O programa possui três módulos: **Transcrição** – destinado à transcrição conservadora do documento, ou seja, sem desenvolver abreviaturas ou marcação de itálico ou negrito, por exemplo; **Edição** – ambiente para a edição do texto, ou seja, para aplicação de etiquetas e definição das camadas de edição que serão aplicadas; e **Morfologia** – voltado à anotação morfossintática automática do texto. Aqui, nos interessa tratar dos dois primeiros.

O processo de edição realizado no eDictor é composto de duas operações básicas: *alteração* de categorias estruturais do texto e *anotação* de categorias estruturais do texto. Para melhor compreensão dessas operações, importa observar como o eDictor transforma um texto simples em um texto estruturado, e quais as estruturas que podem ser modificadas ou anotadas na ferramenta.

De um texto simples produzido na aba Transcrição ou importado de outros processadores (ou seja, qualquer texto em formato .txt) (cf. item 2.0), o eDictor constrói um arquivo em formato XML que organiza as sequências de caracteres em estruturas computacionalmente processáveis. Essa estruturação é realizada automaticamente com base na 'leitura' que o software faz do 'texto' - ou seja, com base nas inferências estruturais realizadas pelo programa a partir da organização dos caracteres. Dessa forma, o eDictor infere que toda sequência de caracteres separada por espaços é uma 'palavra'; que o caractere '.' marca o fim de uma 'sentença'; que todo espaço maior que uma linha é uma divisão de 'parágrafo', etc. O software tenta também inferir os pontos de quebra de página. As inferências realizadas pelo programa sobre quebras de página e separação em sentenças nem sempre são corretas e muitas vezes precisam ser corrigidas pelo usuário.

Idealmente, a leitura gerada tem a seguinte estrutura: no código XML do eDictor, cada parágrafo p contém uma ou mais sentenças s, cada sentença contém uma ou mais palavras w e cada

² Corpus Anotado do Português Tycho Brahe (Universidade Estadual de Campinas); Grupo de Pesquisas Humanidades Digitais (Universidade de São Paulo); Laboratório de História do Português Brasileiro (Universidade Federal do Rio de Janeiro); P.S. – Projeto para o Arquivo Digital de Escrita Quotidiana em Portugal e Espanha na Época Moderna (Universidade de Lisboa); Corpus Eletrônico de Documentos Históricos do Sertão (CE-DOHS) (Universidade Estadual de Feira de Santana); Memória Conquistense (Universidade Estadual do Sudoeste da Bahia); e Projeto do Libolo (Universidade de Macau e Universidade de São Paulo).

palavra contém uma categoria o, significando 'palavra original no texto' (esquematicamente, na figura, todos os caracteres na cor branca fazem parte do texto transcrito; todas as demais cores indicam estrutura construída pelo eDictor). Note-se que os parágrafos, sentenças e palavras são numerados em sequência (automaticamente quando o XML é gerado).

```
▼
 ▼<s id="s 1">
   ▼<w id="1">
      <o>Este</o>
    </w>
   ▼<w id="2">
      <o>é</o>
    </w>
   ▼<w id="3">
      <o>um</o>
    </w>
   ▼<w id="4">
      <o>texto</o>
    </w>
   ▼<w id="5">
      <o>preparado</o>
    </w>
   ▼<w id="6">
      <o>para</o>
    </w>
   ▼<w id="7">
      <o>ser</o>
    </w>
   ▼<w id="8">
      <o>usado</o>
    </w>
   ▼<w id="9">
      <o>como</o>
     </w>
   ▼<w id="10">
      <o>exemplo</o>
    </w>
```

Estrutura XML básica do eDictor: parágrafo, sentença e palavra Importa então esclarecer que as categorias de edição a serem discutidas nesta seção irão interferir em diferentes pontos dessa estrutura, com diferentes qualidades e efeitos - como dito acima, executamos, ao editar, duas operações básicas: alteração de categorias estruturais do texto e anotação de categorias do texto.

Ao acrescentar formas diferentes para cada palavra, estamos alterando a categoria estrutural w. Fundamentalmente, ao inserir edições estamos acrescentando sub-categorias a w, além da categoria original o. A seção **1.1** abaixo mostra essas opções de alteração estrutural, denominadas **editiontypes** no eDictor.

Ao descrevermos aspectos diplomáticos ou paleográficos relevantes, estamos anotando as categorias estruturais **p**, **s**, ou **w**. A seção **1.2** a seguir mostra essas as opções de anotação, denominadas **elementypes** no eDictor.

1.1 Edição de texto e etiquetas

As etiquetas são utilizadas para realizar e sinalizar as intervenções feitas no processo de edição. Elas são criadas de acordo com as necessidades do pesquisador, isso significa que podem atender às diferentes demandas de edição. Servem, por exemplo, para sinalizar as reconstituições por conjectura, as modernizações ortográficas e lexicais, a facilitação de leitura para expectadores leigos e a preparação do texto.

O eDictor vem com duas etiquetas pré-configuradas: **Junção** e **Segmentação**³. Isso significa que, ao se instalar o programa, essas etiquetas já estão disponíveis na lista de "Edição de palavras" por padrão. Ambas têm a função de ajustar as fronteiras entre palavras, permitindo tanto a união de termos que foram separados quanto a separação de termos que aparecem juntos.



Fragmentos do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

No Corpus M.A.P., além das etiquetas pré-configuradas, são utilizadas como etiquetas padrão para a edição dos textos: **Expansão, Grafemática, Atualização** e **Modernização**. A etiqueta **Expansão** é empregada para a abertura de abreviaturas:



³ No eDictor está nomeada como "Ressegmentação".

Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

Destacamos que é possível classificar os diferentes tipos de abreviaturas, permitindo que cada tipo seja associado a uma etiqueta específica. Assim, além de indicar a expansão, a etiqueta também pode ser utilizada para categorizar as abreviaturas, ficando essa decisão a critério do pesquisador. Adiante, são apresentadas as instruções para a inserção de etiquetas.

A etiqueta **Grafemática** é utilizada para uniformização do sistema gráfico do texto, isso inclui padronização de variação caligráfica de sinais diacríticos, por exemplo: [^] e [~] são editados para [~];"sâ" → "sã"; e a uniformização de alógrafos contextuais, por exemplo: "deue" → "deve", "veio" → "vejo", "lustiça" → "Justiça".

Já a etiqueta de Atualização é aplicada para fazer intervenções de acordo com as normas de edição atualizadas, adotadas pelo projeto. Essa etiqueta é aplicada para atualização e normatização ortográfica. Detalhamentos sobre a aplicação das etiquetas são trazidos no item 1.2.



Fragmentos do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

A etiqueta de **Modernização** é utilizada para modernizar o texto, adequando-o às normas de edição modernizada estabelecidas por (inserir referência de Elisa) e adotadas pelo projeto. Essa etiqueta permite intervenções como ajustes na pontuação e atualizações lexicais, garantindo uma versão do texto mais próxima do uso contemporâneo da língua.



Fragmento do processo inquisitorial contra Francisca Luís (ANTT - PT/TT/TSO-IL/028/CX1579/13787)

Além dessas etiquetas padrão para a edição de textos do Corpus M.A.P., outras integram a lista de "Edição de Palavras" do eDictor, adotadas para a documentação do Projeto e apresentadas adiante. Ao criar a lista de etiquetas, é fundamental que o pesquisador observe a ordem em que elas são organizadas, pois existe uma hierarquia a ser respeitada. Caso contrário, podem ocorrer interferências na exportação do texto editado para HTML. A título de exemplo, consideremos a ordenação das etiquetas na figura:

Preferêr	ncias					×
Geral	Edição de Palavras	Elementos do Texto	Morfologia	Metadados XI	ML	
Sigla:	De	scrição:		Adicionar	· >>	
	jun Juncao seg Ressegmentacad exp [Expansao gra Grafematica del Deletado add Adicionado atu Atualizacao mod Modernizacao num Numero					
	ATENÇÃO:	ordene os elemento (do primeiro	os da lista po para o últim	or ordem de ap o).	licação	
	Salvar	Carregar	ОК	Apply	Cancel	

Exemplo de ordenamento das etiquetas na janela "Edição de Palavras"

As etiquetas aqui apresentadas são consideradas padrão porque são aplicáveis a qualquer tipo de edição, seja semidiplomática, atualizada, modernizada ou interpretativa. Sendo necessário adicionar novas etiquetas, é importante posicioná-las após a etiqueta **Grafemática**. Isso se deve ao fato de que as quatro primeiras etiquetas são as únicas utilizadas em edições

semidiplomáticas. Para adicionar novas etiquetas, clique em **Programa** e depois em **Preferências**, conforme indicado:

	E-Dictor [*novo*]		
. ←	Programa Arquivo Editar Exibir Do	umento Edição	Ajuda
- →	Preferências	V A	D 🕅 🗙 🔜 📦
2	Sair Ctrl+Q Propriedades Tipo: Língua: N I S Aplicar alterações Substituir similares	Edição Original: Edição: Forma fonológica:	Análise fonológica Capitular

Caminho para adicionar etiquetas à Edição de Palavras

Ao abrir a janela **Preferências**, clique em **Edição de palavras**. No campo **Sigla**, insira a sigla correspondente à etiqueta que será adicionada, como por exemplo, "mod" para etiqueta de **Modernização**. No campo **Descrição**, insira o nome da etiqueta, depois clique em **Adicionar**:

	Preferê	ncias						×
3	Geral	Edição de Pala	vras	Elementos do Texto	Morfologia	Metadados	XML	
4 ┥	Sigla:	mod	Des	scrição: Modernizacad)	Adicion	ar >>	
		jun Juncao seg Ressegmen exp Expansao gra Grafematica	tacao a	5		(5	
	X	mod Moderniza	cao					
		ATENÇ	ÃO: 0	ordene os elemento (do primeiro p	s da lista po para o últim	or ordem de o).	aplicação	
		Salvar		Carregar	ок	Apply		Cancel

Caminho para adicionar etiquetas à Edição de Palavras

Para mudar a ordem das etiquetas, clique em cima daquela que deseja movimentar e utilize as **setas nas laterais** para colocar na posição que deseja. Também é possível excluir etiquetas, repetindo o processo de seleção e clicando no **botão com "X"**. Por fim, clique em **Apply** e depois em **OK**:

Eargao ae r alavro	S Flemen	tos do Texto	Morfologia	Metadados	XMI	
mod [Descrição:	Modernizaca	D	Adicior	nar >>	
jun Juncao seg Ressegmentac exp Expansao gra Grafematica atu Atualizacao mod Modernizacao	ao			_		
monimoden iizacao						
1						
~						
ATENÇÃO	: ordene (os elemento lo primeiro (s da lista po para o último	r ordem de)).	aplicação	
ATENÇÃO Salvar	: ordene ((Carregar	os elemento do primeiro ; 	s da lista po para o último Οκ	r ordem de)). Apply	aplicação	Cancel
	jun Juncao seg Ressegmentace exp[Expansao gra Grafematica atu Atualizacao mod Modernizacao	jun Juncao seg Ressegmentacao exp[Expansao gra Grafematica atu Atualizacao mod Modernizacao	jun Juncao seg Ressegmentacao exp Expansao gra Grafematica atu Atualizacao mod Modernizacao	jun Juncao seg Ressegmentacao exp[Expansao gra Grafematica atu Atualizacao mod Modernizacao	jun Juncao seg Ressegmentacao exp Expansao gra Grafematica atu Atualizacao mod Modernizacao	jun Juncao seg[Ressegmentacao exp[Expansao gra Grafematica atu Atualizacao mod Modernizacao

Caminho para adicionar etiquetas à Edição de Palavras

O eDictor permite salvar o arquivo de configurações. Isso significa que é possível salvar as preferências de um determinado texto ou até mesmo criar um arquivo de configurações padrão para compartilhar com outros pesquisadores ou para usar diferentes configurações de preferências para textos distintos. O Projeto M.A.P. utiliza um arquivo de configurações padrão como base para as edições, o qual é modificado, isto é, são inseridas etiquetas à medida em que as necessidades de edição se apresentam.

1.1.1 Efeitos estruturais e resumo

Os efeitos estruturais dizem respeito à estrutura organizacional que vai sendo criada em virtude da edição. Cada intervenção realizada gera uma camada extra e ('edição') na palavra estrutural w, e cada camada fica identificada com o atributo do tipo de edição. No exemplo, temos o item original Secretr.o alterado para Secretario > Secretário > secretário , na ordem exp > atu > mod (as camadas se aplicam de baixo para cima no código).



Exemplo de efeito estrutural

Em resumo, as categorias de edição de palavras usadas pela equipe M.A.P. são as seguintes (conforme apresentadas no arquivo de preferências <u>MAP.cfg</u>):

```
editiontypes = jun | Juncao, seg | Ressegmentacao,
exp | Expansão, gra | Grafematica, atu |
Atualização, mod | Modernização
```

1.2 Descrição de elementos de texto

Na edição de texto, também é possível marcar as intervenções em diferentes níveis: **sentença**, **parágrafo**, **seção** e **palavra**. Estas anotações são descritores relativos à estrutura gráfica e intervenções editoriais, e descritores relativos à análise codicológica, diplomática e paleográfica.

Para adicionar um Elemento do Texto, clique em **Programa** e depois em **Preferências**, conforme indicado:

	E-Dictor [*novo*]	
1	Programa Arquivo Editar Exibir Doc	cumento Edição Ajuda
2	Preferências	🕨 🛃 💭 🖾 🗙 🔂 🏓
	Sair Ctrl+Q	
	Transcrição corção hioritología	
	Propriedades	Edição
	Tipo:	Original:
	Língua: N I S	Edição:
	Aplicar alterações	Forma fanalúsica:
	Substituir similares	Porma tonologica;

Caminho para configurar Elementos do Texto

Ao abrir a janela **Preferências**, clique em **Elementos do Texto**. No campo **Elemento**, selecione a camada desejada, como por exemplo, "Palavra". No campo **Tipo**, coloque o nome do descritor que deseja inserir, por exemplo, "ilegível". Depois, clique em **Adicionar**. Por fim, clique em **Apply** e depois em **OK**:



Caminho para configurar Elementos do Texto

Diferente da lista de Edição de Palavras, para a qual é necessário seguir uma ordenação hierárquica, para a lista de Elementos do Texto é indiferente a ordem, uma vez que o eDictor organizará por ordem alfabética.

Ao adicionar descritores que contenham mais de uma palavra, como em **margem esquerda**, recomenda-se a utilização de hífen: **Tipo: margem-esquerda > Adicionar > "Sentença** [margem-esquerda]].

Para aplicar dois descritores em uma única palavra, é necessário criar e adicionar um descritor novo a suas preferências, e para registrar o tipo destes dois descritores pode-se utilizar "+" entre um e outro. Por exemplo, um topônimo que está abreviado a partir de uma letra sobreposta: **Tipo: topônimo+sobrescrito > Adicionar > Palavra |topônimo+sobrescrito| |**.

É possível construir listas de descritores necessários para a sua edição, sendo esta parte da edição bastante flexível. A seguir, apresentamos alguns exemplos de descritores em cada um dos elementos possíveis:

Preferêr	ncias						\times
Geral	Edição de Palavras	Elementos do Texto	Morfologia	Metadados	XML		
Elemer	nto: 🔍 🗸						
Tipo:			para análise	morfológica			
Estilo (CSS (opcional):				Adicio	onar >>	
×	Seção antônimrpereii Seção desconhecido Seção joãopguimara Seção josétcamargo Seção miguelmazeve	ra s do					
	Salvar	Carregar	ОК	Apply		Cancel	

Elementos do Texto - Exemplo 1

Preferêr	ncias							\times
Geral	Edição de Pal	avras	Elementos do Texto	Morfologia	Metadados	XML		
Elemer	nto:	~						
Tipo:			🗌 Ignor	ar para análise	e morfológica			
Estilo (CSS (opcional):					Adici	ionar >>	
×	Parágrafo laber Parágrafo lasse Parágrafo lence Parágrafo lence Parágrafo lence Parágrafo lence Parágrafo loura Parágrafo loura Parágrafo loura Parágrafo lence Parágrafo lence	tura entada ência aminham erramen erramen mento ero-pg ecer ame ign essa	 hento lto lto+remessa orar					
	Salvar	c	arregar	ОК	Apply		Cancel	

Elementos do Texto - Exemplo 2

Preferências			×
Geral Edição de Palavras Elementos do Texto Morfologia Metadados	XML		
Elemento:			
Tipo:			
Estilo CSS (opcional):	Adici	onar >>	
Palavra antropônimos Palavra antropônimos +sóncope Palavra antropônimos +sobrescrito Palavra conjectura Palavra [fragmentado Palavra mancha Palavra polimorfismo-gráfico Palavra polimorfismo-gráfico Palavra sigla Palavra sigla Palavra sincope Palavra sobrescrito Palavra copônimos palavra copônimos +polimorfismo-gráfico			
Salvar Carregar OK Apply		Cancel	

Elementos do Texto - Exemplo 3

Preferêr	ncias						×
Geral	Edição de Palavras	Elementos do Texto	Morfologia	Metadados	XML		
Elemer	nto: 🔍 🗸						
Tipo:		🗌 Ignorar	r para análise	morfológica			
Estilo (CSS (opcional):				Adici	onar >>	
×	Sentença lassinatura Sentença lcarimbo ligi Sentença ldarções l Sentença ldaspedida Sentença ldispositivo Sentença ldispositivo Sentença lexposição Sentença lmargem-di Sentença lmargem-su Sentença lmargem-su Sentença lmargem-su Sentença lposterior liu Sentença lpreâmbulo Sentença lprecação l Sentença lprecação l Sentença lexelositivo Sentença lexelos	 horar ão reita reita reita ferior ferior pperior pporar horar r					
	Salvar	Carregar	ОК	Apply		Cancel	

Elementos do Texto - Exemplo 4

1.2.1. Seção

No elemento **seção**, fazemos anotações referentes àquela seção do texto. O início da seção é indicado por [section], destacado na figura. É possível criar descritores relativos à análise diplomática, como anotar quem é o autor ou a tipologia da seção, conforme exemplificado respectivamente:

[section:miguelmazevedo]	
[p] [s:margem-esquerda] Desacatos ao [s:datações] <u>1795</u> [s:margem-direita] 1 \ [s:margem-esquerda] Santíssimo Sacramento [s:posterior] N 9738 \	
[p] [s] Sumário Contra Antônia \ Maria mulher parda \ moradora na Vila \ do Pitangui Comarca do Sabará \ bispado de Mariana \ de Minas Gerais.	

Seção de Miguel Martins de Azevedo. Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

[section:assentada]
[p] [s] Assentada [s:posterior] 5 [s:margem-direita] 2 \
[p:assentada] [s] Aos vinte e dois dias do mês de Dezembro \ de mil setecentos, e noventa e quatro \ anos nesta Vila de Nossa Senhora do \ Pilar de Pitangui em Casas de moradas do \
Muito Reverendo José Teixeira Camargo \
Vigário da Vara Juiz das Justificações, casamentos, \ Capelas, e Resíduos nesta Vila \
e seu Termo onde Eu Escrivão audiente \
nomeado, e sendo aí apareceram presente \
cujos Nomes, cognomes, Idades \
Pátrias, Moradas, Ofícios, ditos, e Costumes \
são, os que se seguem do que para constar \
Taço este Termo de Assentada \ e Fu O Padre João Pereira Guimarães \
Escrivão do Juízo eclesiástico que \
O Escrevi \

Seção de Assentada. Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

Novamente, a escolha de como aplicar as anotações no nível da seção depende do texto em edição. No caso de uma seção possuir mais de uma tipologia, como por exemplo, uma

assentada e um termo de encerramento na mesma folha, podemos fazer essas anotações em outros níveis, como veremos adiante.

1.2.2. Parágrafo

Como o nome sugere, no elemento **parágrafo** são feitas anotações no nível do parágrafo, no qual aplicamos descritores também relativos à análise diplomática, como por exemplo, encaminhamentos, requerimentos, termos, assentadas, etc.

[p:t-remessa] [s] Termo de remessa \	
[p>:t-remessa] [s] Aos treze dias do mês de Janeiro de \ mil sotecentos, e noventa, e Cinco \ anos nesta Vila de Nossa Senhora \ do Pilar de Pitangui em Casas de \ moradas do Muito Reverendo José Teixeira \ Camargo Vigário da Vara nesta \ Vila, e seu Termo onde Eu Escrivão \ audiente nomeado foi vindo e sendo \ aí faço remessa destes Autos, e ditos \ de Testemunhas neles perguntados \ os ques vão remetidos O Muito Reverendo \ Doutor Antônio de Meireles \ Pereira Rebelo Vigário da vila de \ Caeté, e Comissário do Santo Ofício \	
deste Bispado de Mariana, do que	

Parágrafo onde se inicia Termo de Remessa. Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT -PT/TT/TSO-IL/028/09738)

Aqui também podem ser inseridos descritores como as partes constitutivas de cada documento, por exemplo, dispositivo, expositivo, despedida (valedictio), etc.



Parágrafo onde se inicia o dispositivo. Fragmento do documento peticionário de Caterina Dias de Sene (APESP – 1.1.609/93-1-32)

1.2.3. Sentença

No elemento **sentença**, podemos criar quebras e aberturas de sentenças, conforme necessidade, e utilizar este elemento para descritores paleográficos, como por exemplo, indicar assinaturas ou textos à margem, entre linhas, datações, anotações posteriores ao documento, etc. Podemos também usar descritores como as partes constitutivas do documento.

[section:josétcamargo]
[serem pessoas de crédito, e que falariam verdade, pois \ não tinham motivos de dizer o contrário, e as que juraram \ sobre o crédito das testemunhas, são pessoas qualificadas \ daquele continente e por isso os reputo \ fidedignas. Isto é o que eu entendo na presente \ conjuntura [s:datações] Vila de Pitangui aos]
[p [s:precação] VIGArio da Vara

Descritores anotados no nível da Sentença. Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

Um outro exemplo de descritor que pode ser utilizado tanto no elemento **parágrafo** quanto no elemento **sentença** é o **Reclame**. A escolha da utilização em um ou outro vai depender da estrutura de cada texto.

1.2.4 Tipo de palavra

Das categorias de elementos do texto, a instância da **palavra** é a menor, na qual fazemos anotações paleográficas, codicológicas, de estrutura gráfica e intervenções editoriais, além de adicionar descritores para classificação de palavras. Os descritores padrão para cada palavra, utilizados pelo projeto, são:

Preferências	×
Geral Edição de Palavras Elementos do Texto	Iorfologia Metadados XML
Elemento:	
Tipo:	ara análise morfológica
Estilo CSS (opcional):	Adicionar >>
Palavra antropônimos Palavra apócope Palavra borrado Palavra fragmentado Palavra ilegivel Palavra indix line Palavra mancha Palavra mancha Palavra notas-tironianas Palavra notas-tironianas Palavra solutros Palavra sigla Palavra sigla Palavra sinal-especial Palavra sincope Palavra sobrescrito Palavra topônimos	
Salvar Carregar	K Apply Cancel

Descritores padrão utilizados no Projeto M.A.P.

Assim como na camada das etiquetas e diferente dos últimos três elementos, alguns descritores são considerados padrão em qualquer tipo de edição, por tanto, compõem o arquivo de configurações M.A.P., mas é possível configurar os descritores de acordo com a necessidade do texto e da edição.

Os descritores **topônimos** e **antropônimos** são utilizados para classificar tais ocorrências, tornando possível construir índices para o documento trabalhado e são categorias importantes para a etapa de catalogação de documentos no Portal M.A.P.

Já **sigla**, **apócope**, **síncope**, **sobrescrito**, **sinais especiais**, **notas tironianas** e **outros** são descritores para classificação de diferentes tipos de abreviatura⁴.

Sigla - utilizada quando a palavra é representada pela letra inicial. Estando a palavra no plural, a letra aparece dobrada. Exemplos: "q" → "que"; "VV." → "Vossas"; "SS" → "Senhorias".

Apócope - utilizada quando a palavra é abreviada cortando-se o final. Exemplo: "Guim." \rightarrow "Guimaraens".

Síncope - utilizada quando a palavra é abreviada com elementos do início e do fim da palavra, suprimindo-se o meio da palavra. Exemplo: "Escriam" → "Escrivam".

Sobrescrito - utilizada para palavras que são abreviadas utilizando-se de uma ou mais letras sobrepostas à palavra. Exemplo: "Bisp.^{do}" → "Bispado".

⁴ Classificação feita a partir de FLEXOR, Maria Helena Ochi. Abreviaturas: manuscritos dos séculos XVI ao XIX. 3ª ed. Rio de Janeiro: Arquivo Nacional, 2008. E NUNES, Eduardo Borges. Abreviaturas paleográficas portuguesas. Lisboa: ABRE, 1980.

Os **sinais especiais** são sinais que significam uma sílaba ou ausência de uma letra:



Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)



Fragmento do processo inquisitorial de Domingas Fernandes (ANTT - PT/TT/TSO-IL/028/01281)

As **notas tironianas** são baseadas nas letras do alfabeto maiúsculo romano, sendo sinais utilizados em diferentes posições e que apresentam significados diferentes, de acordo com sua posição⁵:



Fragmentos do processo inquisitorial de Domingas Fernandes (ANTT - PT/TT/TSO-IL/028/01281)

⁵ De acordo com Flexor (2008, p. 12-13).

O descritor **outros** foi criado para categorizar abreviaturas cujas características não se enquadram em nenhuma das categorias anteriores, é o caso de "lanr" → "laneiro" e "Camg." → "Camargo"⁶.

Os descritores **borrado**, **fragmentado**, **manchado**, **rasurado** e **ilegível** são utilizados para anotações que determinam diferentes fragmentações materiais e/ou dificuldades de leitura. O descritor **borrado** é utilizado para sinalizar palavras cuja tinta está borrada no texto, mas que acontece pelo movimento do punho ou do instrumento de escrita:



Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)



Fragmento do documento de Paula Maria Machada (APESP - 1.1.697/24)

O descritor **manchado** é utilizado para sinalizar palavras que se apagaram ou estão parcialmente visíveis por mancha, seja de tinta respingada ou tinta vazada do verso do fólio, por tinta de carimbos ou cola de selos:

⁶ Nota-se que nos dois exemplos, os autores suprimiram tanto letras no meio da palavra, quanto a última letra, fazendo com que esse tipo de abreviatura não seja nem uma síncope nem um apócope. Exemplos retirados do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738).



Fragmentos do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

Em ambos os descritores, podemos fazer a reconstituição do trecho borrado ou manchado por conjectura, como nos exemplos demonstrados, de acordo com as normas de edição e escolhas da editora.

Já o descritor **fragmentado** é utilizado para marcar fragmentações materiais, sejam elas de qualquer natureza, como por exemplo, letras que sumiram por refilamento, palavras desaparecidas por rasgos ou trechos perdidos por ação de insetos, etc. Neste descritor também pode-se optar por reconstruir o texto por conjectura, conforme as normas de edição.

en solo

"Esperamos do seu zelo, e activide. haja # cumprir assim esta nossa detrem@i@ nação que de tudo nos avizara em resposta sua na margem desta. # os nosso Senhor Gde. a VM. Lxa. no Sto. Offo. em # @z@a 15 de Julho de 1795"

> Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

O descritor **rasurado**⁷ é aplicado em trechos que foram rasurados pelo próprio autor do texto, que podem estar totalmente visíveis ou não, assim como pode ser uma palavra escrita por cima de outra palavras:



\$<u>₽</u>"

Fragmento do documento peticionário de Ana Borges (APESP - 1.1.613/93-3-73)

⁷ Chamamos de "rasura" a anulação ou o cancelamento de um elemento por meio de riscos ou de outro tipo de traços contemporâneos à execução do texto (Toledo Neto, 2020, p. 199).

Por fim, o descritor **ilegível** é utilizado para palavras ilegíveis, tanto por fragmentação material quanto por dificuldade de decifração da escrita, mas que não foram reconstituídas por conjectura, seja por impossibilidade ou por escolha.



"G*nho"

Fragmento do Sumário Crime de Antônia Maria de Almeida (ANTT - PT/TT/TSO-IL/028/09738)

1.2.5 Como fazer anotações

Criados os descritores nos quatro elementos do texto, de acordo com as predileções e necessidades de cada editor, segue-se o seguinte passo a passo:

1 - Com o texto transcrito já disponível no módulo de **Edição** do eDictor, clique no elemento ao qual se quer adicionar uma anotação:

1-	[section]
	[p] [s:margem-esquerda] Desacatos ao [s:datações] 1795 [s:margem-direita] 1 \ [s:margem-esquerda] Santíssimo Sacramento [s:posterior] N 9738 \
	[p] [s] Sumário Contra Antônia ∖ Maria mulher parda ∖ moradora na Vila ∖ do Pitangui Comarca do Sabará ∖ bispado de Mariana ∖ de Minas Gerais.

Seção do elemento que receberá anotação

2 - Clique em Propriedades:



Secão do elemento aue receberá anotacão

3 - No campo **Tipo**, selecione o descritor que quer adicionar e clique em **Ok**.

Propried	dades [sc_1]		×
Tipo: Língua:	antônimrpereira desconhecido joãopguimaraes josétcamargo miguelmazevedo	 Formatação Negrito Itálico Sublinhado Cancelar 	
Propried	lades [sc_1]		×
Propriec Tipo: Língua:	ades [sc_1] miguelmazevedo	 Formatação Negrito Itálico Sublinhado 	×

Seção do descritor que será aplicado

Este passo a passo serve para os elementos **Seção, Parágrafo** e **Sentença**. Já para adicionar descritores no nível **Palavra**, é necessário:

1 - Com o texto transcrito e disponível no módulo de **Edição**, clique na palavra a ser adicionado um descritor. No ambiente **Propriedades**, em destaque na figura, selecione o tipo de descritor:

Transcrição Edição Morfologia				
Propriedades	Edição			
Tipo: topônimos ~	Original:	Geraes.	Análise fonológica	Capitular
Língua: NIS	Edição:			>>
Aplicar alterações	Forma fonológica:			
Texto:				
[section:miguelmazevedo]				
[p] [s:margem-esquerda] <mark>Desacatos ao</mark> [s:datações] 1795 [s:margem-direita] 1 \ [s:margem-esquerda] <mark>Santíssimo Sacramento</mark> [s:posterior] N 9738 \				
[p] [s] Sumário Contra Antônia \ Maria mulher parda \ moradora na Vila \				
do Pitangui Comarca do Sabará \				
bispado de Mariana \ de Minas Gerais. Clique sobre a palavra				

Módulo de Edição com destaque no campo Propriedades e na palavra que receberá um descritor 2 - No campo **Tipo**, clique em qual descritor será anotada a palavra escolhida:

Transcrição Edição Morfologia					
Propriedades		Edição			
Tipo: topônimos 🗸		Original:	Geraes.	Análise fonológica	Capitular
Língua: antropônimos antropônimos +síncope antropônimos +sobrescrito	S	Edição:			>>
apócope borrado Subs conjectura		Forma fonológica:]	
Texto: ilegível					
mancha [sectio outros polimorfismo-gráfico					
[p] [s:: rasurado sigla [s:mar sinal-especial	10 (s:da crame	itações] 1795 [s:margem-d ento [s:posterior] N 9738	ireita] 1 \ \		
[p] [s] sobrescrito Maria topônimos mora topônimos +polimorfismo-gráfic					
do Pltangur Cornarca do Sabar bispado de Mariana	á١				
de Minas <mark>Gerais.</mark>					

Seleção do Tipo de descritor

3 - Caso exista mais de uma ocorrência da palavra escolhida no mesmo documento, selecione o item **Substituir similares**, para que o descritor seja aplicado em todas as palavras iguais à selecionada. Depois clique em **Aplicar alterações**.

Transcrição Edição Morfologia						
Propriedades	Edição					
Tipo: topônimos 🗸	Original:	Geraes.	Análise fonológica	Capitular		
Língua: NIS	Edição:			>>		
Aplicar alterações	Forma fonológica:]			
Texto:						
[section:miguelmazevedo]						
[p] [s:margem-esquerda] Desacatos ao [s:d [s:margem-esquerda] Santíssimo Sacram	atações] 1795 [s:margem-d <mark>ento</mark> [s:posterior] N 9738	ireita] 1 \ \				
[p] [s] Sumário Contra Antônia \ Maria mulher parda \ moradora na Vila \ do Pitangui Comarca do Sabará \ bispado de Mariana \						
de Minas Gerais.	de Minas Gerais.					

Aplicação das alterações feitas

1.2.6 Efeitos estruturais e resumo

```
v<s id="s_1" t="margem-esquerda">
v<w id="1" t="sin">
<o>Infr.e</o>
<e t="exp">Informe</e>
</w>
```

Exemplo de efeito estrutural

Vamos observar o efeito estrutural das descrições de elementos de texto descritas acima. Ao contrário do que acontece com as intervenções descritas em **1.1**, as descrições apenas acrescentam anotações à estrutura já existente, na forma de atributos XML aos elementos. No exemplo, a sentença número l está descrita como anotação na margem esquerda (**t="margem** - **equerda**"), e a palavra estrutural (original **Infr.e**, edição de expansão **Informe**) está descrita como contendo uma síncope (**t="sin"**). Todos os descritores apresentados nesta seção são acrescentados dessa maneira aos elementos do XML, sem modificar sua estrutura fundamental.

Em resumo, as categorias de edição de palavras usadas pela equipe M.A.P. são as seguintes (conforme apresentadas no arquivo de preferências <u>MAP.cfg</u>):

```
elementtypes =
Palavra |adição| |, Palavra |apo| |, Palavra |assinatura| |, Palavra
|carimbo|ignorar|, Palavra |conjectura| |, Palavra |exclusão| |, Palavra
|grafismo|ignorar|, Palavra|ilegível| |, Palavra|inserção|ignorar|,
Palavra |out| |, Palavra |selo|ignorar|, Palavra |ses| |, Palavra |sic|
|, Palavra |sig| |, Palavra |sin|, Palavra|tachado| |, Parágrafo |
despacho | , Parágrafo | numero-pagina | ignorar |, Parágrafo |
reclame|ignorar|, Parágrafo | subtitulo| |, Parágrafo | titulo| |,
Sentença | assinatura|ignorar|, Sentença | datação| |, Sentença |
despacho |, Sentença | despedida |, Sentença | direção |, Sentença |
dispositivo |, Sentença | encaminhamento |, Sentença | expositivo |,
Sentença | fecho_complementar | , Sentença | fórmula_de_encerramento | ,
Sentença | identificação | |, Sentença | inserção | |, Sentença | margem-
direita|ignorar|, Sentença | margem-esquerda|ignorar|, Sentença | margem-
inferior | ignorar |, Sentença | margem-superior | ignorar |, Sentença | punho |
|, Sentença | remessa | |, Sentença | requerimento | |
```

1.3 Metadados

Os metadados (os dados sobre os dados) devem ser inseridos à edição eletrônica. Para tanto, fazemos uso da estrutura de metadados oferecida pelo eDictor, em combinação com as categorias catalográficas do Catálogo M.A.P., como mostramos a seguir.

1.3.1 Metadados originais do eDictor

O eDictor oferece uma estrutura muito simples para os metadados, que não pode ser editada; entretanto, o *conteúdo* dos metadados é inteiramente customizável. O XML construído pela

ferramenta apresenta automaticamente um cabeçalho com a estrutura abaixo, preenchido também automaticamente apenas com informações internas de cada arquivo (data de criação, título etc.).

ncias							×
Edição de Palavras	Elementos do 1	Texto	Morfologia	Metadados	XML		
	~	Nome				Adicionar >>	
edition Date edition Editor edition Level original manuscript original manuscript original manuscript original manuscript	Archive Author Catalog name Date Title						
	cias Edição de Palavras edition Date edition Editor edition Level original manuscript original manuscript original manuscript original manuscript	cias Edição de Palavras Elementos do 1 edition Date edition Editor edition Level original manuscript Archive original manuscript Catalog name original manuscript Date original manuscript Title	cias Edição de Palavras Elementos do Texto edition Date edition Editor edition Level original manuscript Archive original manuscript Catalog name original manuscript Date original manuscript Title	cias Edição de Palavras Elementos do Texto Morfologia edition Date edition Editor edition Level original manuscript Archive original manuscript Catalog name original manuscript Date original manuscript Title	cias Edição de Palavras Elementos do Texto Morfologia Metadados edition Date edition Date edition Level original manuscript Archive original manuscript Catalog name original manuscript Date original manuscript Title	cias Edição de Palavras Elementos do Texto Morfologia Metadados XML w Nome: edition Date edition Editor edition Level original manuscript Archive original manuscript Author original manuscript Date original manuscript Title	cias Edição de Palavras Elementos do Texto Morfologia Metadados XML v Nome: Adicionar >> edition Date edition Editor edition Level original manuscript Archive original manuscript Catalog name original manuscript Date original manuscript Title

Interface de inserção de Metadados no eDictor

```
<?xml-stylesheet href="" type="text/xsl"?>
 <head id="GP">
    <metadata generation="edictor_internal">
        <n>Document Name</n>
        \langle v \rangle GP \langle v \rangle
      </meta>
        <n>XML generated by</n>
        <v>E-Dictor-v1.0.b010</v>
      </meta>
        <n>Last Saved Date</n>
        <v>31.08.2023</v>
      <meta>
        <n>Word Count</n>
        <v>2670</v>
      </meta>
    </metadata>
```

Note-se a estrutura básica dos metadados: [metadata [meta [n v]]] ; note-se também que o elemento metadata recebe um atributo generation ; neste caso, o valor do atributo para os dados gerados automaticamente é edictor_internal .

Essa estrutura é fixa e incontornável; entretanto, o eDictor permite que o usuário acrescente quantas camadas de **metadata** quiser, e que as classifique como quiser no atributo **generation**.

Fazemos uso desse recurso no Corpus, como mostrado em **1.3.2**.

Alternativamente, se há muitos dados de Metadados a serem inseridos, isso pode ser feito diretamente no XML depois que ele for gerado pelo eDictor (e a ferramenta lerá normalmente essas informações em uma próxima vez).

1.3.2 Uso customizado dos metadados do eDictor no Corpus M.A.P.

Como o eDictor é a base de todas as edições filológicas feitas no M.A.P., precisamos manter nos documentos XML editados a estrutura de metadados oferecida pelo eDictor. Partindo dessa contingência, desenvolvemos uma adaptação na qual todas as categorias incluídas na descrição catalográfica do M.A.P. se tornam compatíveis com a estrutura do eDictor, sendo o mapeamento entre as duas estruturas facilitado. As imagens abaixo permitem comparar a estrutura adaptada do cabeçalho de um XML gerado pelo eDictor à estrutura do XML original do Catálogo M.A.P.

```
?xml version='1.0' encoding='utf-8'?>
<head id="FL">
 <metadata generation="MAP">
     <n>codigo MAP</n>
     <v>0031</v>
   <n>codigo</n>
   <v>97FL</v>
   <n>nome modernizado</n>
   <v>Francisca Luís</v>
   <n>grafia conservadora</n>
   <v>Frca Luis; Frca luis; francisca Luis</v>
   <n>trecho nomeacao</n>
   <v>Francisca Luís, mulher preta, forra,
   crioula, da cidade do Porto, casada com
   Domingos Soares, homem pardo, remendão,
   ausente, do qual não tem novas se é vivo
   se morto, vendedeira, moradora nesta
   cidade</v>
```

Metadados eDictor: geração M.A.P.

Como dito acima, o eDictor permite a inserção de quantas categorias metadata sejam desejadas, desde que preservada a estrutura [metadata [meta [n v]]:

```
<metadata generation= ('...')>
    <meta>
        <n>('...')</n>
        <v>('...')</v>
        </meta>
</metadata>
```

Em nossa adaptação, inserimos uma categoria metadata com o atributo generation="MAP", e usamos a estrutura n ('nome') e v ('valor') para guardar o nome e valor das categorias catalográficas M.A.P. do documento.

Esse sistema é ilustrado ainda pelas imagens ao lado, que mostram a visualização do arquivo XML de base para a instância catalográfica 0031, Francisca Luís.

xml version='1.0' encoding='utf-8'?
<document></document>
<head id="FL"></head>
MAP metadata
<pre><metadata generation="MAP"></metadata></pre>
<pre><!--edictor metadata--></pre>
<pre>(metadata generation="edictor internal")</pre>
chodys
<pre></pre>
title="Denúncia contra Francisca Luís"
author="Manuel Francisco" vear="1592">
$\langle sc id = sc 1 \rangle$
<pre><n id="n 1" t="margin ton left"></n></pre>
$\langle s id = "s 1" \rangle$
<pre><w id="1" t="antropônimo"></w></pre>
<pre><o>frca</o></pre>
<pre><e t="exp">francisca</e></pre>
<pre><e t="atu">Francisca</e></pre>
<pre></pre>
<pre><w id="2"></w></pre>
<o>Luis</o>
<pre><e t="mod">Luís</e></pre>
<w id="3"></w>
<o>negra</o>

Pode-se ver a estrutura XML com o detalhe do cabeçalho referente aos dados catalográficos M.A.P. e o detalhe do cabeçalho referente aos dados gerados internamente pelo eDictor; e depois a estrutura do documento editado, com o cabeçalho em colapso.

A partir dessa junção entre o cabeçalho XML do eDictor e o antigo XML do Catálogo M.A.P. surge o formato básico do Portal M.A.P., no qual cada instância catalogada corresponde apenas a um

documento XML, que inclui os dados catalográficos M.A.P. (na forma dos Metadados eDictor) e, se for o caso, a edição filológica do documento.

Como tratamos com maior detalhe a seguir, cada um desses documentos XML (bem como o conjunto dos XMLs) pode gerar diferentes visualizações HTML com diferentes finalidades. No momento, temos a visualização de Fichas catalográficas individuais, Edições (semidiplomáticas, modernizadas, etc.) e uma visualização com todas as Fichas reunidas (com ligações para o conjunto documental de cada instância).

Importa observar que, no futuro próximo, essa estrutura de metadatos será modificada, em particular tendo em vista a remodelação do eDictor em um formato de XML compatível com o TEI (Text Encoding Initiative). Nesse sentido, pensamos que a opção por tomar a estrutura atual dos metadados do eDictor como base para o M.A.P. é muito apropriada, pois o novo eDictor 2.0 conterá mecanismos simples de importação do XML da versão 1.0. Assim, ao serem importados para o eDictor 2.0 nossos arquivos automaticamente terão seus metadados convertidos em um formato estandardizado.

1.4 Apresentação final das edições

1.4.1 Visualizações das camadas de edição

O eDictor permite a geração de diferentes tipos de visualização dos documentos editados, a partir do arquivo XML de base. Por padrão, essas visualizações são: Texto simples, Hipertexto e Hipertexto com área de fac-símile. Cada uma dessas visualizações pode mostrar quantas e quais camadas de intervenção editorial sejam desejadas.

Assim, é possível gerar por exemplo arquivos .html com área de fac-símile e o texto em versão original; ou com texto até a camada grafemática; ou com texto até a camada modernizada, etc.

Programa	Arquivo Editar Exibir	Documento Edição	Ajuda
7 J	Novo		🕞 🕅 🗙 🖪 🃦
Transcrição Propriedad	Abrir Abrir imagem Abrir recente	CTRL+O	
Língua:	Salvar Salvar como	Ctrl+S CTRL+SHIFT+S	
	Reverter	F5 :	
Texto:	Importar Exportar Fechar	> CTRL+W	Texto Léxico de edições
[p] [s:marg	em-esquerda] Desacatos a	0 [s:datações] <u>1795</u> [s:	Sentenças etiquetadas

Caminho para exportação de texto no eDictor

Exportar texto	×
Nível de edição:	Texto original 🛛 🗸
Tipo de arquivo:	Texto (TXT) 🗸
Opções Exportar forma fonológ	Texto (TXT) Hipertexto (HTML) Hipertexto c/ área de fac-simile (HTML)
Exportar uma sentença Realizar quebras de linh Não exportar nomes de	por linha a elementos
Ok	Cancelar

Caminho para exportação de texto no eDictor

No M.A.P., por padrão, oferecemos os textos na versão Texto simples contendo até a última camada de edição (atualização ou modernização), e na versão Hipertexto com área de facsímile em pelo menos duas camadas: edição diplomática e semidiplomática. Se aplicada, também a modernizada.

1.4.2 Aparência e estilos

O estilo ou aparência do HTML gerado automaticamente pelo eDictor é adaptado no Corpus M.A.P. para melhor visualização das categorias de edição utilizadas, e também para que se adaptem mais organicamente à estética geral do site. Temos então a seguinte diferença, obtida apenas pela aplicação de estilos em cascada (Cascading Style Sheets) ao HTML básico do eDictor (da versão texto com área de fac-símile):

HTML original do eDictor (cabeçalho e fólio 1r do documento 0031) HTML customizado M.A.P. (cabeçalho e fólio 1r do documento 0031)



As opções atuais de estilo do M.A.P. estão no arquivo

https://map.prp.usp.br/Styles/edictor_2025.css. Note-se que esse estilo se aplica apenas às categorias de html específicas do eDictor (ou seja, para a aparência externa completa, é preciso que ele se combine com o estilo geral do M.A.P, que está em https://map.prp.usp.br/Styles/MAP_2025.css).

Importa notar que para aplicar este estilo a um HTML gerado pelo eDictor, basta ativá-lo no cabeçalho do documento. Os estilos incidem diretamente sobre o código do eDictor, não sendo necessário fazer mudanças na estrutura HTML (a não ser que se deseje acrescentar cabeçalhos, menus, etc.).

2. Normas de edição

As Normas de edição seguidas no Corpus M.A.P. procuram conjugar os modos de edição tradicionais com o método de edição digital em camadas. Para melhor compreensão do conceito de edição em camadas, lembramos que o texto estruturado pelo eDictor toma a palavra gráfica como menor unidade da estrutura, mas permite que cada 'palavra' seja organizada em sub-itens. Como visto em **1. Processo de edição de texto**, no arquivo de base todas as intervenções e a forma original da palavra permanecem disponíveis; além disso, as intervenções são nomeadas por tipos (como visto, intervenção de *expansão, atualização ortográfica*, etc.). Repetimos aqui o exemplo apresentado em **1. Processo de edição de texto**:



No exemplo, cada intervenção realizada na palavra estrutural w, originalmente composta apenas pela camada original o (Secretr.o) está codificada como uma camada extra e ('edição') identificada com o atributo do tipo de edição (exp 'expansão', atu 'atualização ortográfica', mod 'modernização'). Chamamos de 'camadas' de edição essas diferentes

intervenções aplicadas, e que podem ser recuperadas por tipos na apresentação final, sendo possível selecionar quais dessas camadas serão visíveis. No exemplo, uma apresentação conservadora mostrará **Secretr.o** e uma apresentação modernizada mostrará **secretário.**, como ilustram as imagens abaixo.



Informe o senhor coronel secretário. São Paulo, 12 de fevereiro de 1820.

Trecho da apresentação diplomática de 0007

Trecho da apresentação modernizada de 0007

Nas subseções que seguem, listamos as camadas de intervenção na ordem das menos invasivas para as mais invasivas, o que se traduzirá em apresentações finais do documento nos moldes das formas de edição filológica tradicional (*diplomática, semidiplomática, modernizada*).

2.0 Etapa de Transcrição

A transcrição antecede as edições a serem realizadas no eDictor. Esta etapa pode ser realizada no próprio software ou em um arquivo no formato .txt que pode ser importado para a ferramenta. Para que a edição possa ser realizada sem grandes problemas, é importante seguir algumas normas de transcrição. Fundamentalmente, importa lembrar que o texto transcrito será transformado em um texto estruturado em XML pelo eDictor, e que ao estruturar o texto, o software vai separar a sequência de caracteres em unidades - sendo a menor unidade da estrutura (o *token*) a palavra gráfica e o sinais de pontuação. A palavra gráfica é definida como uma sequência de caracteres separada por um espaço, e os sinais de pontuação tomados como tokens são: *ponto final, vírgula, dois pontos, ponto e vírgula, ponto de exclamação, ponto* *e vírgula, parênteses, colchetes e chaves.* As figuras (a) e (b) ilustram essa questão: a segunda figura mostra o resultado da criação de um XML a partir do texto simples da primeira. Observese que algumas sequências digitadas sem espaço (ex.: **[colchetes] , !exclamação!**) foram interpretadas pelo software como tokens separados.

Interpretados como tokens e marcas de novas sentenças
.ponto. ?interrogação? !exclamação!
Interpretados como tokens
,vírgula, ;ponto e vírgula; (parêntesis) [colchetes] {chaves} "aspas" 'aspas simples'
Interpretados como caracteres normais
<pre><parêntesis angular=""> @arroba@ #tralha# \$cifrão\$ %porcentagem% *asterisco* =igual= +adição+ -hífentravessão_ /barra/ \barra invertida\ barra vertical </parêntesis></pre>
Texto digitado no módulo Transcrição do eDictor.
[p] [s] Interpretados como tokens e marcas de novas sentenças
[p] [s] . ponto . [s] ? interrogação ? [s] ! exclamação !
[p] [s] Interpretados como tokens

[p] [s] , vírgula , ; ponto e vírgula ; (parêntesis) [colchetes] { chaves } " aspas " ' aspas simples '

[p] [s] Interpretados como caracteres normais

[p] [s] <parêntesis angular> @arroba@ #tralha# \$cifrão\$ %porcentagem% *asterisco* =igual= +adição+ -hífen-_travessão_ /barra/ \barra invertida\ |barra vertical|

Resultado da geração de XML do texto acima.

Assim, é importante evitar na transcrição o uso de sinais tokenizáveis (por exemplo, colchetes para indicar conjecturas ou comentários paleográficos).

As normas para transcrição em texto simples usadas no Corpus estão descritas no Quadro 1:

1	A transcrição é realizada sem marcas de formatação (como negrito ou itálico) e sem colchetes.
2	A transcrição é sempre justalinear. Linhas em branco são reproduzidas assim como ocorrem no modelo, mas não são sinalizadas.
3	Maiúsculas e minúsculas são reproduzidas fielmente. Abreviaturas, sinais de pontuação, sinais de acentuação e paragrafação são também transcritos fielmente. A separação vocabular é reproduzida conservando-se a fronteira de palavras que ocorre no modelo.
4	Caracteres alfabéticos e numéricos são transcritos em caracteres redondos atuais.
5	Caracteres de <i>leitura duvidosa</i> por dificuldade de decifração da escrita são transcritos por conjectura e <mark>podem ser sinalizados entre travessões: _travessão_</mark> . Posteriormente, na etapa de edição no eDictor, tais caracteres podem ser anotados como <i>conjectura</i> , ou de acordo com a preferência da editora.

Quadro 1 - Normas de Transcrição

6	Caracteres de <i>leitura impossível</i> por dificuldade de decifração da escrita são transcritos como asteriscos. Usa-se um asterisco – * – para letra, e dois asteriscos – ** – para palavra.
7	Elementos <i>inexistentes</i> por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta são transcritos como uma tralha: #.
8	Elementos <i>parcialmente visíveis</i> por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta são transcritos por conjectura e podem ser sinalizados entre tralhas: #arroba#. Posteriormente, na etapa de edição no eDictor, tais caracteres podem ser anotados como <i>conjectura</i> , ou de acordo com a preferência da editora.
9	Caracteres <i>suprimidos</i> são transcritos como um hífen. Ex.: <i>-uprimido</i> (no caso de s suprimido). <i>Rasuras</i> , sejam palavras riscadas ou palavras escritas em cima de palavras, são transcritas a partir do que se lê entre sinal de igual: =igual=. Posteriormente, na etapa de edição no eDictor, tais rasuras podem ser discriminadas e anotadas de acordo com as preferências da editora.
10	Elementos <i>repetidos</i> são transcritos assim como aparecem no modelo e podem ser sinalizados entre barras verticais: barra vertical . Posteriormente, na etapa de edição no eDictor, tais elementos podem ser anotados como <i>repetições</i> ou de acordo com a preferência da editora.
	Caracteres nas entrelinhas ou margens são transcritos normalmente na posição que se encontram no texto. Posteriormente, na etapa de edição no eDictor, tais elementos Quadro elaborado pela equipe do Projeto M.A.P. podem ser discriminados e anotados de acordo com as preferências da editora.
11	
12	Mudanças de punho não são sinalizadas na etapa da transcrição. Posteriormente, na etapa de edição no eDictor, tais elementos podem ser discriminados e anotados de acordo com as preferências da editora.
13	Anotações posteriores são transcritas normalmente e na posição original que se encontram. Posteriormente, na etapa de edição no eDictor, tais elementos podem ser discriminados e anotados de acordo com as preferências da editora.

Quadro elaborado pela equipe do Projeto M.A.P.

2.1. Camada diplomática

Tradicionalmente, a "Edição diplomática" é considerada a edição mais conservadora, onde se exerce menos interferência no texto. Neste tipo de edição, o texto transcrito deve estar o mais próximo possível do texto original, mantendo-se todas as suas características originais, como fronteiras de palavras, abreviaturas, alografias, etc. A intervenção editorial busca ser mínima, limitando-se às indicações de problemas de leitura (deterioração material, por exemplo).

Para trabalhar essa camada em nosso método digital, seguimos as seguintes normas que podem ser ajustadas para mais conservadoras de acordo com a preferência da editora:

1	A camada diplomática mantém o texto justalinear.
2	Caracteres alfabéticos e numéricos são mantidos em caracteres redondos atuais.
3	Maiúsculas e minúsculas são mantidas assim como aparecem no modelo.
4	Abreviaturas são mantidas assim como aparecem no modelo.
5	Sinais de pontuação são mantidos assim como aparecem no modelo.
6	A separação vocabular é mantida de acordo com a fronteira realizada no modelo.
7	A paragrafação é mantida assim como aparece no modelo.
8	Os sinais de acentuação são uniformizados apenas quanto a sua variação caligráfica: ^ para ~ ; nos ditongos, mantém-se fielmente a posição dos sinais que se depositem sobre um caractere apenas: aõ.
9	Caracteres de <i>leitura duvidosa</i> por dificuldade de decifração da escrita transcritos por conjectura entre travessões (_travessão_) devem ser anotados no nível da palavra, sinalizando-se que se trata de uma transcrição por conjectura.
10	Caracteres de <i>leitura impossível</i> por dificuldade de decifração da escrita são transcritos como asterisco *. Usa-se um asterisco para letra e dois para palavra.
11	Elementos <i>inexistentes</i> por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta são transcritos como uma tralha #.
12	Elementos <i>parcialmente</i> visíveis por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta transcritos entre arrobas (@arroba@) devem ser anotados no nível da palavra, sinalizando-se que se trata de uma reconstituição por conjectura. Nesta etapa também pode-se usar o mesmo recurso para discriminar o tipo de fragmentação.
17	Caracteres suprimidos transcritos com um hífen (- <i>uprimido)</i> devem ser anotados no nível da palavra, sinalizando-se que se trata de uma supressão.
18	Rasuras transcritas entre sinal de igual (=igual=) devem ser anotadas no nível da palavra, sinalizando-se que se trata de uma rasura. Nesta etapa também pode-se usar o mesmo recurso para discriminar o tipo de rasura.
19	Elementos repetidos, transcritos assim como aparecem no modelo e sinalizados entre barras verticais, podem ser anotados no nível da palavra, sinalizando-se que se tratam de repetições.

Quadro 2 - Normas da Camada Diplomática

19	Caracteres nas entrelinhas ou margens são transcritos normalmente na posição que se encontram no texto, e sinalizados através de anotações nos diferentes níveis que o software dispõe (parágrafo e sentença), de acordo com as especificidades de cada texto/editora.
20	Mudanças de punho não sinalizadas na etapa da transcrição, são anotadas no elemento transcrito, através de anotações nos diferentes níveis que o software dispõe (seção , parágrafo e sentença).
21	Anotações posteriores transcritas normalmente e na posição original que se encontram são sinalizadas através de anotações nos diferentes níveis que o software dispõe (parágrafo e sentença), de acordo com as especificidades de cada texto/editora.

Quadro elaborado pela equipe do Projeto M.A.P.

Unindo essa explicação ao que foi apresentado em **1. Processo de edição**, a camada diplomática não inclui nenhuma das intervenções listadas em **1.1 Edição de texto** – ou seja: nesta camada, não há intervenções no sentido da adição de versões editadas ao texto. A camada contém, entretanto, diversas das anotações para descrição de elementos de texto listadas em **1.2 Descrição de elementos de texto**. Em resumo, nesta camada são aplicados todos os descritores referentes aos aspectos paleográficos e diplomáticos considerados relevantes pela editora.

2.2 Camada semidiplomática

Tradicionalmente, a "**Edição semidiplomática**" apresenta a maioria das características conservadoras da edição diplomática, mas aqui já são realizadas algumas intervenções com o intuito de que a edição se torne mais acessível a um público menos especializado. É possível, através deste tipo de edição, manter características do texto original, mas realizar-se algumas intervenções, de modo que a leitura deste tipo de edição seja mais fácil - como, por exemplo, a normalização de alógrafos e a expansão de abreviaturas.

A escolha entre uma edição diplomática ou semidiplomática depende dos objetivos de cada editora - sendo fundamental lembrar sempre que, em nossa metodologia, um arquivo trabalhado até a camada semidiplomática pode sempre gerar uma visualização que se limite às intervenções mais conservadoras da camada diplomática.

Para trabalhar a camada semidiplomática em nosso método digital, acrescentamos à camada diplomática anotações relativas a esse tipo de intervenções mais substantivas; em nosso sistema, apenas quatro normas desta camada são diferentes das normas da camada diplomática, sendo elas:

1	A camada semidiplomática mantém o texto justalinear.
2	Caracteres alfabéticos e numéricos são mantidos em caracteres redondos atuais.
3	Maiúsculas e minúsculas são mantidas assim como aparecem no modelo.
4	As abreviaturas são expandidas. A grafia de partes ausentes, quando não existem outras ocorrências no texto, segue a ortografia atual. Caso existam ocorrências no texto, segue-se a ortografia já presente no texto original. Utiliza-se a etiquetas Expansão. .
5	Sinais de pontuação são mantidos assim como aparecem no modelo.
6	As fronteiras entre palavras são modernizadas conforme o modelo de separação vocabular atual. Utiliza-se as etiquetas Junção e Ressegmentação .
7	A paragrafação é mantida assim como aparece no modelo.
8	Os alógrafos são uniformizados: I é transcrito como J; u para v. Por exemplo: "deue" → "deve", "veio" → "vejo", "lustiça" → "Justiça". Utiliza-se a etiqueta Grafemática .
9	Os sinais de acentuação são uniformizados quanto a sua variação caligráfica: ^ para ~ ; nos ditongos, mantém-se fielmente a posição dos sinais que se depositem sobre um caractere apenas: aõ. Utiliza-se a etiqueta Grafemática .
10	Caracteres de <i>leitura duvidosa</i> por dificuldade de decifração da escrita transcritos por conjectura entre travessões (_travessão_) podem ser editados através de etiquetas ou anotados no nível da palavra, para sinalizar que se trata de uma transcrição por conjectura e remover o travessão.
11	Caracteres de <i>leitura impossível</i> por dificuldade de decifração da escrita são transcritos como asterisco *. Usa-se um asterisco para letra e dois para palavra.
12	Elementos <i>inexistentes</i> por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta são transcritos como uma tralha #.
13	Elementos <i>parcialmente</i> visíveis por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta transcritos entre arrobas (@arroba@) podem ser editados através de etiquetas ou anotados no nível da palavra, para sinalizar que se trata de uma reconstituição por conjectura e remover o arroba. Nesta etapa também pode-se usar o mesmo recurso para discriminar o tipo de fragmentação.
14	Caracteres suprimidos transcritos com um hífen (- <i>uprimido)</i> podem ser editados através de etiquetas ou anotados no nível da palavra, para sinalizar que o elemento suprimido foi reconstituído por conjectura e remover o hífen.
15	Rasuras transcritas entre sinal de igual (=igual=) podem ser editadas através de etiquetas ou anotados no nível da palavra, para sinalizar que se trata de uma rasura e remover o igual. Nesta etapa também pode-se usar os mesmos recursos para discriminar o tipo de rasura.
16	Elementos repetidos, transcritos assim como aparecem no modelo e sinalizados entre barras verticais, podem ser anotados no nível da palavra, sinalizando-se que se tratam de repetições.

Quadro 3 - Normas da Camada Semidiplomática

1/ Mu	udanças de punho não sinalizadas na etapa da transcrição, são anotadas no elemento
trai	anscrito, através de anotações nos diferentes níveis que o software dispõe (seção, parágrafo e
ser	entença).
18 And	notações posteriores transcritas normalmente e na posição original que se encontram são
sind	nalizadas através de anotações nos diferentes níveis que o software dispõe (parágrafo e
ser	e ntença), de acordo com as especificidades de cada texto/editora.

Quadro elaborado pela equipe do Projeto M.A.P.

Unindo essa explicação ao que foi apresentado em **1. Processo de edição**, a camada semidiplomática inclui diversas das intervenções listadas em **1.1 Edição de texto** - ou seja: nesta camada, ao contrário da camada diplomática, já há intervenções no sentido da adição de versões editadas ao texto. Os tipos de edição listados em **1.1** correspondentes à camada semidiplomática são: *Junção, Segmentação, Expansão, Grafemática* (ou seja, respectivamente: e t="jun", e t="seg", e t="exp" e t="gra").

Quanto às classes de descritores de elementos de texto listadas em **1.2 Descrição de elementos de texto**, fazem parte da camada semidiplomática os mesmos descritores correspondentes à camada diplomática (tais sejam, *borrado, fragmentado, manchado*, etc), e mais os descritores correspondentes aos tipos de edição - nomeadamente, aqueles relativos às abreviaturas: *sigla, apócope, síncope, sobrescrito, sinais especiais, notas tironianas*, aplicados à edição de tipo *Expansão*. Em resumo, nesta camada são aplicadas as intervenções editoriais mais conservadoras (expansão de abreviaturas, normalização grafemática) e sua tipologia (sendo herdados, da camada diplomática, todos os descritores referentes aos aspectos paleográficos e diplomáticos considerados relevantes pela editora).

2.3 Camada atualizada

A **Camada atualizada** pode ser compreendida como o meio termo entre as camadas mais conservadoras e as camadas mais modernizadas de interferência no texto. Além da editora atualizar as fronteiras de palavras (juntando e separando os elementos de acordo com o uso atual), expandir abreviaturas e realizar a uniformização grafemática, opera, também, a atualização ortográfica – mas não a atualização lexical.

A atualização da grafia e dos sinais gráficos é feita, na **Camada atualizada**, de acordo com a ortografia vigente: "annos", "huã" e "naó", por exemplo, têm a grafia atualizada para "anos", "uma" e "não". Por outro lado, não são realizadas alterações no nível lexical dos termos editados: a palavra "couza", por exemplo, é modificada para "cousa", e não para "coisa"; "pareçeo" é modificado para "pareceu", embora possa ter sido utilizado em seu contexto original com o sentido de "apareceu". Do mesmo modo, a editora não realiza a atualização do emprego de minúsculas e maiúsculas. Modifica-se, por exemplo, "s. paullo" para "são paulo", e não para "São Paulo" ; "vex^a." para "vossa excelência", e não para "Vossa Excelência"; e "Prezensa" para

"Presença", não para "presença". Por fim, também não se alteram os sinais de pontuação. Abaixo, apresentamos de forma resumida tal proposta de normas para uma **Camada atualizada**:

1	A camada atualizada mantém o texto justalinear.						
2	Caracteres alfabéticos e numéricos são mantidos em caracteres redondos atuais.						
3	Maiúsculas e minúsculas são mantidas assim como aparecem no modelo.						
4	As abreviaturas são expandidas. A grafia de partes ausentes segue a ortografia atual. Utiliza-se a etiqueta Expansão . Para tal operação, recomenda-se a consulta a dicionários de abreviaturas (cf. Nunes, 1981, e Flexor, 2008;2020). Ex.: " p 'tanto" é modificado para " portanto "; " Ill^{mo}. eEx^{mo}. S '." é modificado para " Ilustríssimo e Excelentíssimo Senhor ".						
5	Sinais de pontuação são mantidos assim como aparecem no modelo.						
6	As fronteiras entre palavras são modernizadas conforme o modelo de separação vocabular atual. Utiliza-se as etiquetas Junção e Ressegmentação . Ex.: " p ' tanto" é modificado para " p'tanto "; " seprostra " é modificado para " se prostra "; " se de- ve " (separado por quebra de linha) é modificado para " se deve ".						
7	A paragrafação é mantida assim como aparece no modelo.						
8	Os alógrafos contextuais, como I/J e u/v, são uniformizados. Ex.: " deue " é modificado para " deve "; " veio " é modificado para " vejo "; " lustiça " é modificado para " Justiça ". Utiliza-se a etiqueta Grafemática .						
9	Os sinais de acentuação são uniformizados quanto a sua variação caligráfica: ^ para ~ ; nos ditongos, atualiza-se a posição dos sinais que se depositem sobre um ou mais caracteres. Ex.: " - aõ " é modificado para " -ão ". Utiliza-se a etiqueta Grafemática .						
10	Caracteres de <i>leitura duvidosa</i> por dificuldade de decifração da escrita, transcritos por conjectura entre travessões (_travessão_), podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de uma conjectura e, assim, remover o travessão. O resultado de tal operação pode variar de acordo com a preferência da editora.						
11	Caracteres de <i>leitura impossível</i> por dificuldade de decifração da escrita, transcritos como asterisco *, podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de um trecho ilegível. Usa-se um asterisco para letra e dois, para palavra. O resultado de tal operação pode variar de acordo com a preferência da editora.						
12	Elementos <i>inexistentes</i> por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta, transcritos como uma tralha #, podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de um trecho ilegível.						
13	Elementos <i>parcialmente</i> visíveis por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta, transcritos entre tralhas (#arroba#), podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de uma conjectura e, assim, remover o arroba. Nesta etapa, também pode-se usar o mesmo recurso para discriminar o tipo de fragmentação.						

Quadro 4 - Normas da Camada Atualizada

14	Caracteres <i>suprimidos</i> , transcritos com um hífen (<i>-uprimido</i>), podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que o elemento suprimido foi
	reconstituido por conjectura e, assim, remover o mien.
15	<i>Rasuras</i> transcritas entre sinal de igual (=igual=) podem ser editadas através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de uma rasura e, assim, remover o sinal de igual. Nesta etapa também pode-se usar os mesmos recursos para discriminar o tipo de rasura.
16	Caracteres nas entrelinhas, ou nas margens, transcritos normalmente na posição em que se encontram no texto, são sinalizados através de anotações nos diferentes níveis que o <i>software</i> dispõe (parágrafo e sentença), de acordo com as especificidades de cada texto e as preferências de cada editora.
17	Mudanças de punho não sinalizadas na etapa da transcrição são sinalizadas através de etiquetas de anotação nos diferentes níveis que o <i>software</i> dispõe (seção, parágrafo e sentença).
18	Anotações posteriores, transcritas normalmente e na posição original que se encontram no modelo, são sinalizadas através de anotações nos diferentes níveis que o <i>software</i> dispõe (parágrafo e sentença), de acordo com as especificidades de cada texto e as preferências de cada editora.
19	A grafia de nomes comuns, topônimos e formas de tratamento é atualizada de acordo com o uso atual, excetuando-se a modificação de maiúsculas e minúsculas. Ex.: "annos " é modificado para "anos "; "huã " é modificado para "uma "; e "naó " é modificado para "não "; "Prezensa " é modificado para "excelência "; "s. paullo " é modificado para "são paulo" ; "vexª." é modificado para "vossa excelência "; "baya de todos os fanctos" é modificado para "bahia de todos os santos" . A grafia de antropônimos também pode ser padronizada de acordo com sua ocorrência atual mais comum. Ex.: "Jzabel " e "Izabel " podem ser modificados para "Isabel "; "Feliçiana ", "Filiciana " e "feliciana ".
20	A atualização não é feita no nível do léxico. Também não são corrigidas, por exemplo, as variações em concordância que são entendidas atualmente como erros. Ex.: "Coufas" é modificado para "Cousas" ; "pareçeo " é modificado para " pareceu "; "Ill^{mo}. e Ex^{mos}. Senhores " é modificado para "Ilustríssimo e Excelentíssimos Senhores ".

Quadro elaborado pela equipe do Projeto M.A.P.

Unindo essa explicação ao que foi apresentado em **1. Processo de edição**, a camada atualizada, além de herdar todos os elementos já aplicados nas camadas diplomática e semidiplomática (expansão de abreviaturas e suas especificações; normalização grafemática; descrição diplomática), corresponde a uma primeira camada menos conservadora, incluindo agora as intervenções do tipo *Atualização ortográfica* (e t="atu").

2.4 Camada modernizada

Tradicionalmente, a "**Edição modernizada**" pode ser considerada como o maior nível de interferência editorial no texto. Nela, estão compreendidas não somente todas as operações realizadas na **Camada atualizada**, acima melhor descrita, como também atualizações lexicais, padronizações de maiúsculas e minúsculas e, por fim, alterações na pontuação. Não se

recomenda, porém, alterações sintáticas profundas, nem a modernização extrema de antropônimos, a ponto do texto original ficar irreconhecível. A seguir, apresentamos uma proposta resumida de normas para **Camada modernizada**:

Quadro	5 -	Normas	da	Camada	Modern	izada
Quuuio	0	nonnus	uu	Cumuuu	MOUEIII	izuuu

1	A camada modernizada mantém o texto justalinear.
2	Caracteres alfabéticos e numéricos são mantidos em caracteres redondos atuais.
3	Maiúsculas e minúsculas são modificadas de acordo com o uso atual.
4	As abreviaturas são expandidas. A grafia de partes ausentes segue a ortografia atual. Utiliza-se a etiqueta Expansão .Para tal operação, recomenda-se a consulta a dicionários de abreviaturas (cf. Nunes, 1981, e Flexor, 2008;2020). Ex.: " p 'tanto" é modificado para " portanto "; " Ill^{mo}. eEx^{mo}. S '." é modificado para " Ilustríssimo e Excelentíssimo Senhor ".
5	Sinais de pontuação são alterados, quando necessário, visando apenas a facilitação da compreensão do sentido do texto. Ex: a pontuação do trecho " ella sup^e. he senhora, epossuidora do sittio, que erdou de seo Pai " é retirada, sendo este modificado para " ela suplicante é senhora e possuidora do sítio que herdou de seu pai ".
6	As fronteiras entre palavras são modernizadas conforme o modelo de separação vocabular atual. Utiliza-se as etiquetas Junção e Ressegmentação . Ex.: " p^r tanto " é modificado para " p'tanto "; " seprostra " é modificado para " se prostra "; " se de- ve " (separado por quebra de linha) é modificado para " se deve ".
7	A paragrafação é mantida assim como aparece no modelo.
8	Os alógrafos contextuais, como I/J e u/v, são uniformizados. Ex.: " deue " é modificado para " deve "; " veio " é modificado para " vejo "; " lustiça " é modificado para " Justiça ". Utiliza-se a etiqueta Grafemática .
9	Os sinais de acentuação são uniformizados quanto a sua variação caligráfica: ^ para ~ ; nos ditongos, atualiza-se a posição dos sinais que se depositem sobre um ou mais caracteres. Ex.: " -aõ " é modificado para " -ão ". Utiliza-se a etiqueta Grafemática .
10	Caracteres de <i>leitura duvidosa</i> por dificuldade de decifração da escrita, transcritos por conjectura entre travessões (_travessão_). podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de uma conjectura e, assim, remover o travessão. O resultado de tal operação pode variar de acordo com a preferência da editora.
11	Caracteres de <i>leitura impossível</i> por dificuldade de decifração da escrita, transcritos como asterisco *, podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de um trecho ilegível. Usa-se um asterisco para letra e dois, para palavra. O resultado de tal operação pode variar de acordo com a preferência da editora.
12	Elementos <i>inexistentes</i> por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta, transcritos como uma tralha #, podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de um trecho ilegível.
13	Elementos <i>parcialmente</i> visíveis por corte, rasgo, mancha, descoloração ou qualquer dano do suporte material ou da tinta, transcritos entre arrobas (@arroba@), podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de uma conjectura e, assim,

	remover o arroba. Nesta etapa, também pode-se usar o mesmo recurso para discriminar o tipo de fragmentação.
14	Caracteres suprimidos, transcritos com um hífen (- <i>uprimido</i>), podem ser editados através de etiquetas, ou anotados no nível da palavra, para sinalizar que o elemento suprimido foi reconstituído por conjectura e, assim, remover o hífen.
15	<i>Rasuras</i> transcritas entre sinal de igual (=igual=) podem ser editadas através de etiquetas, ou anotados no nível da palavra, para sinalizar que se trata de uma rasura e, assim, remover o sinal de igual. Nesta etapa também pode-se usar os mesmos recursos para discriminar o tipo de rasura.
16	Caracteres nas entrelinhas, ou nas margens, transcritos normalmente na posição em que se encontram no texto, são sinalizados através de anotações nos diferentes níveis que o <i>software</i> dispõe (parágrafo e sentença), de acordo com as especificidades de cada texto e as preferências de cada editora.
17	Mudanças de punho não sinalizadas na etapa da transcrição são sinalizadas através de etiquetas de anotação nos diferentes níveis que o <i>software</i> dispõe (seção, parágrafo e sentença).
18	Anotações posteriores, transcritas normalmente e na posição original que se encontram no modelo, são sinalizadas através de anotações nos diferentes níveis que o <i>software</i> dispõe (parágrafo e sentença), de acordo com as especificidades de cada texto e as preferências de cada editora.
19	A grafia de nomes comuns, topônimos e formas de tratamento é atualizada de acordo com o uso atual. Ex.: "annos " é modificado para "anos "; "huã " é modificado para "uma "; e "naó " é modificado para "não "; "Prezensa " é modificado para "Presença "; "s. paullo " é modificado para "são paulo "; "vexª." é modificado para "vossa excelência "; "baya de todos os fanctos" é modificado para "bahia de todos os santos" . A grafia de antropônimos também pode ser padronizada de acordo com sua ocorrência atual mais comum. Ex.: "Jzabel " e "Izabel " podem ser modificados para "Isabel "; "Feliçiana ", "Filiciana " e "felicianna " podem ser modificados para "Feliciana ".
20	A atualização é feita também no nível do léxico. Realiza-se, entre outras operações, a correção de variações em concordância entendidas atualmente como erros. Ex.: "Couzas" é modificado para "coisas" ; "pareçeo" é modificado para "apareceu" ; "Ill^{mo}. e Ex^{mos}. Senhores " é modificado para "Ilustríssimos e excelentíssimos Senhores ".

Quadro elaborado pela equipe do Projeto M.A.P.

Unindo essa explicação ao que foi apresentado em **1. Processo de edição**, a camada modernizada, além de herdar todos os elementos já aplicados nas camadas diplomática e semidiplomática (expansão de abreviaturas e suas especificações; normalização grafemática; descrição diplomática) e da camada atualizada (atualização ortográfica), equivale à camada menos conservadora de todas, incluindo agora as intervenções do tipo *Modernização* (e t="mod").

2.5 Versões para visualização

Como apresentado brevemente em **1.4 Apresentação final**, o sistema de edição filológica no eDictor permite a geração de diferentes versões para visualização a partir de um mesmo arquivo de hiperedição em camadas. Atualmente, no Corpus M.A.P., oferecemos versões de leitura correspondentes às camadas *diplomática*, *semidiplomática*, *atualizada* e *modernizada*, com as características abaixo exemplificadas e resumidas.

Aos quinze dias do mes de dezembro de mil e quinhentos e nouenta e dous annos nesta cidade do faluador Capitanja da bahia de

Trecho da apresentação diplomática de 0031

Aos quinze dias do mês de dezembro de mil e quinhentos e noventa e dois anos nesta cidade do Salvador Capitania da Bahia de

Trecho da apresentação modernizada de 0031

Pedem AV.Exa. seja servdo., atentas as razoēs expos= tas, pelo Amor de Deos, advertir o vexame emq. as suptes. estaõ postas, ecom as man-os levantadas, esperaõ huã sau[da] vel providencia. ERM Trecho da apresentação diplomática de 0091 Pedem A Vossa Excelência seja servido, atentas as razões expostas, pelo Amor de Deus, advertir o vexame em que as suplicantes estão postas, e com as mãos levantadas, esperam uma saudável providência. Espera Receber Mercê Trecho da apresentação atualizada de 0091 Pedem a Vossa Excelência seja servido, atentas as razões expostas, pelo amor de Deus, advertir o vexame em que as suplicantes estão postas, e com as mãos levantadas, esperam uma saudável providência.

Espera receber mercê

Trecho da apresentação modernizada de 0091

4. Como referenciar

Ao utilizar os textos e as informações reunidas no *Corpus* M.A.P., é necessário referenciar o banco eletrônico e a(s) edição(ões) utilizada(s).

4.1 O Corpus M.A.P.

Para referenciar o corpus, basta copiar e colar a seguinte referência:

PAIXÃO DE SOUSA, M. C.; MONTE, V. M. (Orgs.). Projeto M.A.P. – Mulheres na América Portuguesa. Disponível em: < https://map.prp.usp.br/ >. Acesso em: dia, mês e ano.

4.2 As edições eletrônicas

Para referenciar um documento, siga o modelo apresentado, fazendo as alterações pertinentes, inserindo o nome da editora e o título do documento:

SOBRENOME DA EDITORA, iniciais do nome (conferir metadados). Documento. In: Paixão de Sousa, M. C.; Monte, V. M. (Orgs.). Projeto M.A.P. – Mulheres na América Portuguesa. Disponível em: < https://map.prp.usp.br/ >. Acesso em: dia, mês e ano.



tutorial de instalação, customização e uso básico Este tutorial foi desenvolvido para auxiliar novas pessoas usuárias do eDictor. Por ser introdutório, são apresentadas orientações para instalação, customização e uso inicial. Detalhamentos sobre as configurações da ferramenta são apresentados no *Manual eDictor 1.0 beta 10,* disponível na página do programa: https://manualedictor.wordpress.com/edictor/.

1. Download e instalação

Acesse o link https://manualedictor.wordpress.com/download/ e faça o download.



Figura 1 - Página para baixar o programa

Depois do download concluído, localize onde o arquivo foi armazenado em seu computador e execute o programa. Para isso, clique no arquivo baixado e siga as instruções de instalação que serão apresentadas.

Como carregar o arquivo de configurações M.A.P.

O conceito fundamental do eDictor é oferecer um ambiente de edição filológica que seja ao mesmo tempo fortemente controlado e amplamente customizável. Assim, a ferramenta em estado 'cru', imediatamente depois da instalação, apresenta apenas três categorias básicas de anotação; todo o resto deve ser personalizado pelo usuário. Essa customização é feita no menu Preferências do software, conforme apresentado em 2. As personalizações podem ser salvas, exportadas e importadas por outros usuários.

Assim, a fim de garantir um padrão de edição dos textos e agilizar o processo de edição, depois de instalar o programa, as editoras deverão carregar no eDictor o arquivo de configurações

(MAP.cfg) com as preferências previamente estabelecidas, correspondentes aos documentos que serão editados.

Para carregar o arquivo de configurações, clique em Programa > Preferências > Carregar:

	E-Dictor [*novo*]		
1	Programa Arquivo Editar Exibir Doo	umento Edição	Ajuda
2	Preferências	1	
	Sair Ctrl+Q	E	
	Halbulyao colyao hortologia		
	Propriedades	Edição	
	Tipo:	Original:	Análise fonológica Capitular
	Língua: N I S	Edição:	>>>
	Aplicar alterações	Forma fonológica:	
	Substituir similares		

Fiaura 2 - Caminho para acessar a opcão Preferências

Geral	Edição de Palavras	Elementos do Texto	Morfologia	Metadados	XML	
Aparêr	ncia					
Lingua	gem: Portugué	s (brasileiro) 🛛 🗸				
√ Real	izar quebra de linha i	na edição				
Segura	inça					
✓ Salv ✓ Grav	ar cópia anterior (cor /ações automáticas e	n extensão "~") m cópia temporária oc	ulta			
		3				
		3				
		3				
		3				

Fiaura 3 - Carreaando o arauivo de confiaurações

Ao clicar em **Carregar**, o programa abrirá uma janela para o explorador de arquivos do computador. Localize o arquivo de configurações que deseja carregar e clique em **Abrir**:



Figura 4 - Localizando o arquivo de configurações

Feito isso, clique em **Apply** > **Ok**, conforme apresentado em 5, e feche a caixa de diálogo:



Fiaura 5 - Localizando o arauivo de confiaurações

2. Módulo de transcrição

O módulo de transcrição é o primeiro módulo disponível no eDictor. Para iniciar a transcrição, carregue o fac-símile do documento clicando no ícone imagem:

Programa	Arquivo	Editar	Exibir	Documento	Edição	Ajuda	
	3	2	-	e .	E	DDX	
Transcrição	Edicão N	Iorfologia					
Te	xto	ion rologic					Fac-sím
\bigcirc \Box						^	

Figura 6 - Ícone para abrir o fac-símile no eDictor

Ao clicar no ícone, o programa abrirá uma janela para o explorador de arquivos do computador. Selecione a imagem desejada e clique em **Abrir**:



Figura 7 - Explorador de arquivos para abrir o fac-símile desejado

Depois de abrir a imagem desejada, observe que à esquerda ficará a área de transcrição e à direita o fac-símile, ambos destacados em 8:

💵 E-Dictor ("novo") Programa Arquivo Editar Exibir Documento Edição Ajuda 📷 🚟 🙈 🙈 🐟 🖝 🕎 🖓 🖓 📄 🕅	×
	Thicke queude Sallet "13-64 Thicke queude Sallet "13-64 led bys corrange . au it worder gatani innensmanne & suite of corr sub parm airde deat ordiner if our manile par sed on the ordiner of our out the part he onits trick our or our altrice, he person a quelque sense the deater succes a the our the pare of a stater to the fact of the other being a . Part by to be the origin a mailer succes have a alsone i bade a mailer succes have to alsone i bade a mailer succes have to alsone i bade a mailer succes have to the set of the to the day a mailer bades a alsone i bade a mailer based and the sense a to the pare desgin a set to the to the or to the pare desgin a set
Rvel de edição a exibiri 🔤 🚾	🏓 🔶 🖬 🗤 🛊 🏟

Figura 8 - Módulo de transcrição com o fac-símile carregado

Se o documento a ser editado já estiver transcrito em formato de texto, não há necessidade de inserir a imagem. Apenas copie e cole o texto na área de transcrição.

Depois de o texto devidamente transcrito/colado, gere o arquivo em XML clicando em **Documento > Gerar XML**:

E-Dictor ("novo")	1	– 🗆 🗙
Programa Arquivo Editar Exibir	Documento Edição Ajuda	
🚃 🚐 🏡 🏡 📥	Gener XMI CTRL+ALT+G	1
	Etiquetagem automática	
Transcrição Edição Morfologia	Metadados CTRL+ALT+M	
IEK00	Navegação	csme
		minila quinda Sallate Maninho Jost haje concernes a que se ancinto; partania interessmente de celta a con mà prese ainde deste un rite à preside que ses one deme presente de preside que he omitte tanto ma one de preside side. Une sale que que omitte heu a de e que fue oracite tente com o que actines, dis presente tente com o que actines, dis

Figura 9 - Caminho para gerar o XML no eDictor

Ao gerar o XML, o programa automaticamente redirecionará a editora para o módulo de edição.

3 Módulo de edição

No módulo de edição é possível fazer a **edição de palavras** e incluir os **metadados**, detalhados a seguir.

Para fazer a edição de palavras, clique sobre a **palavra** e depois clique no campo **Edição**:



Figura 10 - Como editar palavras

Selecione a etiqueta adequada. Ao clicar sobre a etiqueta, a palavra selecionada aparecerá no campo ao lado direito, destacado na Figura 11, como 3:

E-Dictor [C/Users\prisc\OneDrive\Área	de Trabalho\DOUTO	RADO - Priscile - etue	ilizada\Acervo (CAFET)\XM	II. (cartas 105-	136]\112-FMS	- 🗆 X
rograma Arquino Edicar Exileir Do	Cumento Edição /	J 💭 🗙	0 🏓			
ranscrição Edição Morfologia						
Inguia: N 1 S Aplicar alterações	Edição: Forma fonológica:	porem) padronização	Análise fonológica	Capitular	Lists de edições	Burrouge both
Substituir similares Texto:			3		incidents.	Chemical and
(section) (c) (s sectore) Minha querida Sallete ((c) (s) - 1º aninho - \ (c) (s) Você hoje comemora o seu 1º imensamente de estar al com vocês ainda desta vez não é possivel , (s) (aninho ; gostaria \ a, <u>porém</u> \ Queria que \					~

Figura 11 - Campo para editar palavras

Depois de fazer as alterações necessárias, clique no botão ">>" para aplicar as alterações realizadas e confira a lista de edições:

F-Dictor (C-Ulters),princ\OneDrive\A rograma Arquivo Esitar Exibir D	rea de Trabalho),DOUTORAD locumento Edição Ajud	O - Priscila - atualizi b	ada/,Acenvo (CAFET)/,XM	IL (cartas 105-	136),112-FMS — 🗆 🕄
🗺 📁 🤐 🏤 🦘 (iansoigão Edição Morfologia		S X .	0 🏓		5
Propriedades	Edição		7	-	Lista de edições
Lingua: N I S	Edição:		Analise tonologica	Capitular >>	porem padrontração
Aplicar alterações	Forma fonológica:			4	Remover Remover tudo
Texto:				-	
perior p perior Minha querida Sallets p p=10 aninho - \ p p Você hoje comemora o seu	(panin) 4 - 7 - 62 \ 1º aninho ; gostaria \				
imensamente de estar aí com voc ainda desta vez não é possível . p	és , parém \ Queria que \				

Figura 12 - Botão para aplicar alterações e lista de interferências feitas na palavra

Para corrigir alguma eventual inconsistência na edição, utilize os botões: **Remover**, para remover uma das etiquetas; e **Remover tudo**, para remover todas as etiquetas aplicadas.

Após conferir a lista de edições, salve a edição clicando em Aplicar alterações.