



**UNIVERSIDADE ESTADUAL DE FEIRA DE SANTANA**

Autorizada pelo Decreto Federal nº 77.496 de 27/04/76  
Recredenciamento pelo Decreto nº 17.228 de 25/11/2016



**PRÓ-REITORIA DE PESQUISA E PÓS-GRADUAÇÃO**  
COORDENAÇÃO DE INICIAÇÃO CIENTÍFICA

## **XXIV SEMINÁRIO DE INICIAÇÃO CIENTÍFICA DA UEFS SEMANA NACIONAL DE CIÊNCIA E TECNOLOGIA - 2020**

### **EDIÇÃO XML E METADADOS DE DOCUMENTOS HISTÓRICOS**

**Allana Rios de Oliveira<sup>1</sup>; Zenaide de Oliveira Novais Carneiro**

1. Bolsista PIBIC/FAPESB, Graduando em Letras com Inglês, Universidade Estadual de Feira de Santana, e-mail:

[allanarios02@gmail.com](mailto:allanarios02@gmail.com)

2. Orientador, Departamento de Letras e Artes, Universidade Estadual de Feira de Santana, e-mail:

[zenaide.novais@gmail.com](mailto:zenaide.novais@gmail.com)

**PALAVRAS-CHAVE:** Linguística Histórica; Linguística de Corpus; eDictor.

### **INTRODUÇÃO**

A reconstrução da história do português brasileiro, segundo Mattos e Silva (2002), nunca recobriria uma realidade homogênea, ou seja, a língua portuguesa no Brasil apresenta muitas variações, não somente em locais diferentes, mas também em tempos diferentes. E isso se mostra muito no processo de edição de documentos antigos. Com base nisso, o trabalho feito durante a bolsa teve como objetivos editar os documentos em versão eletrônica na linguagem XML e acrescentar os metadados dos escreventes dos documentos que eram testamentos, certidões, declarações e inventários, datados nos séculos XVII e XVIII e escritos na cidade de Cachoeira, disponíveis no repositório do PHPB (Projeto Para a História do Português Brasileiro). Para os estudos linguísticos e de outras áreas, é imprescindível que essas edições sejam o mais fiéis possível aos documentos originais pois, dessa maneira, os estudiosos que farão uso desse material, terão uma análise mais embasada e poderão trazer resultados mais confiáveis. Além da facilidade do acesso e da leitura desses materiais através do site da plataforma. E é nessa perspectiva que o trabalho se mostra importante.

### **MATERIAL E MÉTODOS OU METODOLOGIA (ou equivalente)**

O trabalho foi feito através do aplicativo eDictor, Paixão, Kepler e Faria (2009), que fornece as ferramentas necessárias para que o preenchimento dos metadados e as edições em formato XML aconteça. Através de um sistema que tem etiquetas, cada uma com sua função, o editor pode modernizar a escrita da edição semidiplomática, ou seja, a transcrição do documento original, e assim facilitar a leitura do documento. Dessa forma, o editor tem acesso à edição semidiplomática em formato DOCX e a coloca no aplicativo de edição, na seção “Transcrição”. E, transformando o arquivo em XML, a modernização da escrita dessa versão eletrônica começa e a inserção das informações nos metadados também. Após esses procedimentos, os arquivos são salvos no computador em pastas, contendo nelas o arquivo da edição semidiplomática e o da edição em XML.

### **RESULTADOS E/OU DISCUSSÃO (ou Análise e discussão dos resultados)**

Os resultados propostos no plano de trabalho envolviam a publicação dessas edições no site do projeto de pesquisa CE-DOHS. Por conta da pandemia do novo Coronavírus e do isolamento social, esse trabalho foi feito de forma remota, assim como as orientações. Assim, não consegui ter acesso aos fac-símiles dos documentos, a publicação dessas edições no site do projeto foram impedidas, porém esse conteúdo será providenciado assim que for possível. Mas, mesmo assim, o trabalho foi feito remotamente, pois a ferramenta é facilmente encontrada para download pela internet. Todas as edições em XML (figura 1) e os metadados (figura 2) foram acrescentados, com as informações sobre os escreventes e sobre os documentos encontradas em Carneiro e Almeida (pg. 8-11).

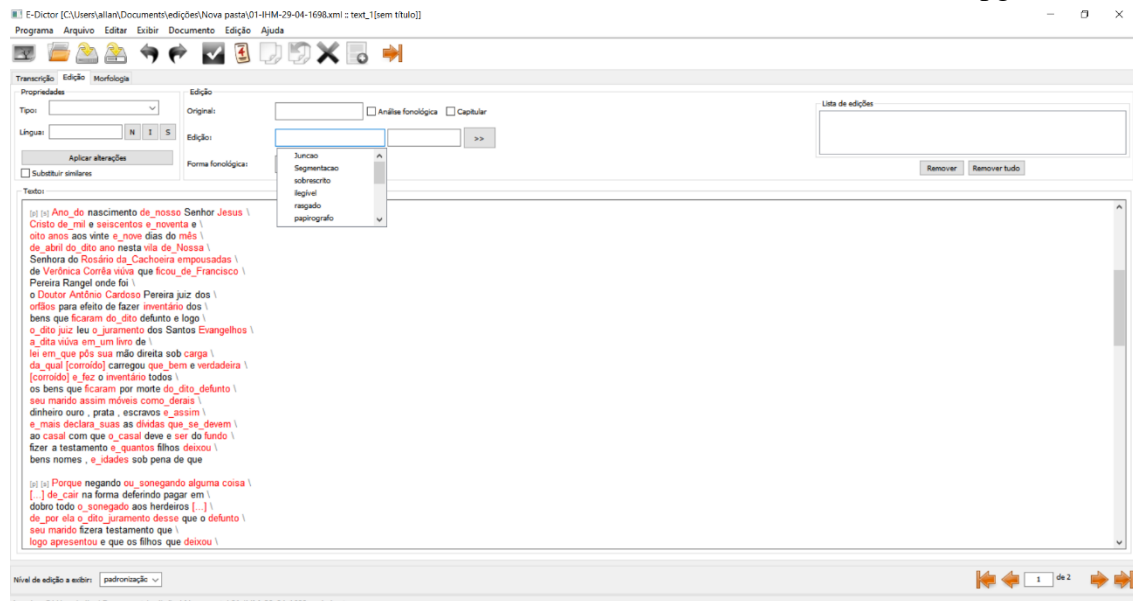


Figura 1

Ficha catalográfica	
Dados do documento	
acervo	CAV
autor	Jozé da Silva Lisboa
conteúdo	Felicita o amigo pelo despacho do serviço militar e informa sobre o lugar de oficial.
destinatário	Manoel Ignacio da Cunha e Menezes (visconde do Rio Vermelho)
genero do documento	carta
local	Rio de Janeiro
relacao social	Amigo
fonte	Instituto Geográfico e Histórico da Bahia, Arquivo Público da Bahia e Arquivo da Casa do Sertão/UEFS
nascido(N)/radicado(R)	Salvador, BA.
referencia	Carneiro, Zenaide de Oliveira Novais. Cartas Brasileiras (1808-1904): um estudo lingüístico-filológico. Tese de Doutorado, Unicamp-- Campinas, SP: 2005, volume 2, tomo 1.
data do documento	16 de setembro de 1809
data de nascimento	16 de julho de 1756
genero do autor	Masculino
Processamento	
edição semi-diplomática	Zenaide de Oliveira Novais Carneiro
revisão da edição semi-diplomática	Zenaide de Oliveira Novais Carneiro
edição XML	Cynthia Yano
1ª revisão da edição XML	Zenaide de Oliveira Novais Carneiro
revisão final da edição XML	Em andamento
número de palavras	214

Figura 2

Durante o processo de edição e da percepção dos detalhes que esse trabalho demandam, é possível notar a importância de cada etapa da edição e das etiquetas, pois essa forma facilita que o leitor perceba os processos pelos quais aquelas palavras passaram para

chegar até a escrita modernizada e também estudar essas mudanças que, no site do projeto, se localizam no Léxico de Edições (figura 3).

Carta pessoal 01-ARAB-13-12-1829

|| Autor: Antonio Rodriguez de Araujo Basto  
 || Destinatário: Senhor Manoel Ignacio da Cunha e Menezes  
 || Data: 13 de dezembro de 1829  
 || [Lista dos itens editados](#) (por ordem alfabética)  
 || [ver ficha catalográfica e outras versões disponíveis](#)

Item Original	Edições	Fonologia
acha rá	seg: achará, jun: achará,	
amo	exp: amigo,	
Ámo	exp: Amigo, sob: o,	
Ánto	exp: Antonio, sob: o,	
Áro	exp: Araujo,	
cro	exp: criado, sob: o,	
cumprímtos	exp: cumprimentos, sob: tos,	
deo	pad: deu,	
derigi= dos	seg: derigidos, jun: derigi=dos, pad: dirigidos,	
deV	seg: de Vossa, exp: de Vossa,	
Dezejando	pad: Desejando,	
eMenes	seg: e Menezes, exp: e Menezes, sob: es,	
eSenhor	seg: e Senhor,	
Exa	exp: Excelência, sob: a,	
Exa	exp: Excelência, sob: a,	
Exa	exp: Excelencia, pad: Excelência, sob: a,	
Exa	exp: Excelencia, pad: Excelência, sob: a,	
Exmo	exp: Excelentissimo, pad: Excelentissimo, sob: mo,	
Exmos	exp: Excelentissimos, pad: Excelentissimos, sob: mos,	
Familia	pad: Família,	
Familia	pad: Família,	
igualmte	exp: igualmente, sob: te,	
Ilmo	exp: Illustrissimo, pad: Ilustrissimo, sob: mo,	

Figura 3

Dessa forma, assim que disponível, esse trabalho se mostra importante para estudos da linguística, direito e outras áreas, facilitando o acesso desse conteúdo a diversas pessoas, além de agregar mais materiais ao banco de dados do CEDOHS. Além disso, ele também disponibilizou mais materiais nos quais podem ser notadas as variações da língua portuguesa conforme o tempo, sendo mais uma evidência de que a mesma não é homogênea.

## REFERÊNCIAS

- CARNEIRO, Zenaide. (org.) **Cartas Brasileiras (1809-2000)**: coletânea de fontes para o estudo do português brasileiro. Vol. 1 (CD-ROM 1, 2 e 3). Feira de Santana: UEFS, 2011;
- MATTOS E SILVA, Rosa Virgínia. **Para a história do português culto e popular brasileiro: sugestões para uma pauta de pesquisa**. In: Alkmim, Tânia M. Para a história do português brasileiro: novos estudos. São Paulo. Humanitas/FFCHL/USP:FAPESP, Vol. III, p. 443-464, 2002.
- Paixão de Sousa, M. C.; Kepler, F. N.; Faria, P. **E-dictor: Novas perspectivas na codificação e edição de corpora de textos históricos**. In: VIII Encontro de Linguística de Corpus, 2009, Rio de Janeiro. Resumos, 2009.
- CARNEIRO, Zenaide de Oliveira Novais de; ALMEIDA, Norma Lúcia Fernandes de. **Notícias sobre o banco de textos de Feira de Santana e Cachoeira**. Texto do III PHPB-Campinas.