PNLD Digital

Relatório de testes para conversão de ebooks utilizando Calibre

Sumário

1

1. Introdução 1

2. Ferramenta de conversão 1

3. Utilização do Calibre 1

3.1 Instalação do *software* 1

3.2 Adicionando uma nova obra 2

3.3 Conversão das obras. 3

3.4 Utilizando a CLI 7

4. Formatos utilizados 7

4.1 Configuração de saída do formato PDF 8

4.2 Validação de saída do formato EPUB 8

5. Realização dos testes 9

5.1 Livro Modelo 1 - Obra simples em arquivo único HTML. 9

5.1.1 Preparação da obra 10

5.1.2 Objetivo 10

5.1.3 Apresentação da obra original - HTML 10

5.1.4 Conversão para EPUB 11

5.1.5 Conversão para MOBI e AZW3 13

5.1.6 Conversão para PDF 15

5.1.7 Conclusão: Livro Modelo 1 19

5.2 Livro Modelo 2 - Obra com metadados e capa (Arquivo OPF) 20

5.2.1 Preparação da obra 20

5.2.2 Objetivo 20

5.2.3 Apresentação da obra original - HTML 21

5.2.4 Validação do formato EPUB 21

5.2.5 Observação dos metadados 22

5.2.6 Observação da Tabela de Conteúdo (TOC) 23

5.2.7 Observação da Capa 25

5.2.8 Conclusão: Livro Modelo 2 27

5.3 Livro Modelo 3 - Obra com múltiplos arquivos e estilos. 27
5.3.1 Preparação da obra 27
5.3.2 Objetivo 28
5.3.3 Apresentação da obra original - HTML 28
5.3.4 Conversão para EPUB 31
5.3.5 Conversão para MOBI e AZW3 35
5.3.6 Conversão para PDF 39
5.3.7 Conclusão: Livro Modelo 3 43
5.4 Livro Modelo 4 - Obra com tags de acessibilidade 43
5.4.1 Preparação da obra 44
5.4.2 Objetivo 49
5.4.3 Teste de leitura e navegação 49
5.4.4 Validação do PDF 52
5.4.5 Conclusão: Livro Modelo 4 55
6. Conclusão 56
7. Referências 59
8. Anexos 60
8.1 Exemplo arquivo toc.ncx. 60
8.2 Exemplo arquivo content.opf 62
8.3 Exemplo arquivo index.xhtml 64
8.4 Exemplo arquivo de conteúdo content.xhtml 66

1. Introdução

Para o desenvolvimento do projeto PNLD Digital foi apontado como mais adequado, a partir de vários estudos, o formato HTML5 como o padrão para o recebimento das obras digitais. O formato HTML5 é bastante difundido para o desenvolvimento e marcação visual de páginas na internet e também serve como base para alguns formatos de livros digitais, como o EPub.

Este documento visa ao registro dos testes de conversão entre alguns formatos de livros digitais a partir de uma versão da obra disponibilizada em HTML5 realizados pela equipe da RNP com o intuito de garantir que os usuários do PNLD Digital tenham a liberdade de visualizar o conteúdo disponível nos mais diversos formatos e equipamentos.

2. Ferramenta de conversão

De acordo com as definições previamente construídas, todos os testes foram realizados utilizando a ferramenta de conversão multiformato disponibilizado pelo *software* Calibre (https://calibre-ebook.com/), que disponibiliza uma interface gráfica e uma aplicação de linha de comando que pode ser utilizada para processos automatizados e integrações com *softwares* de terceiros.

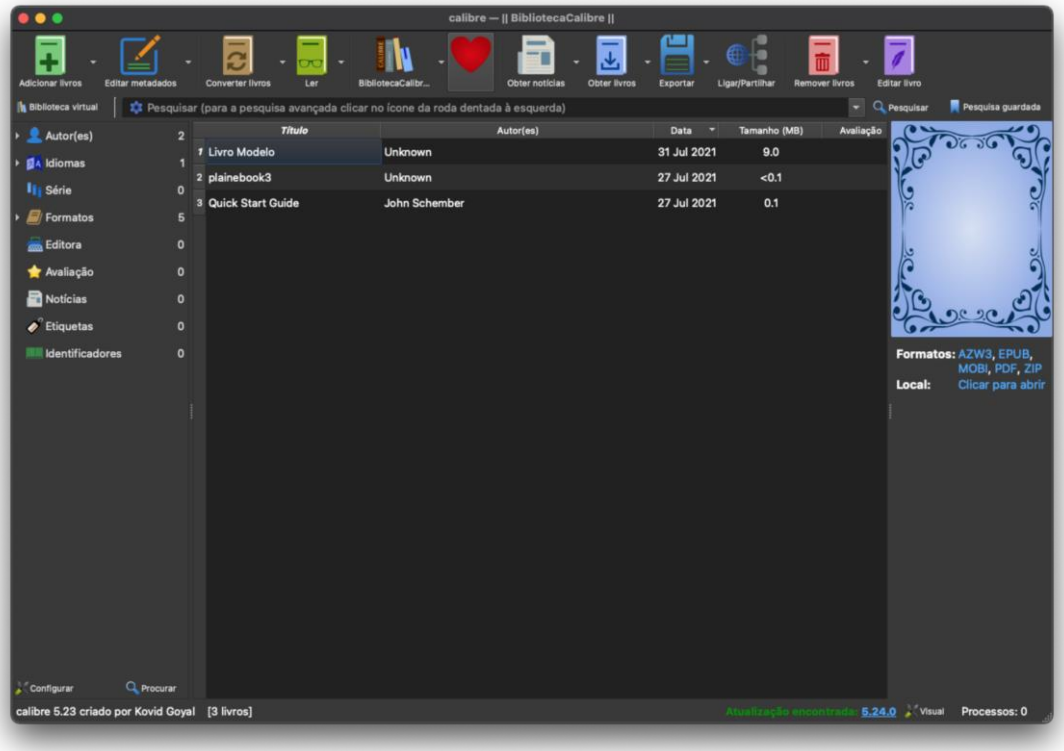
Os testes foram realizados utilizando a ferramenta visual do Calibre, versão 5.23.0, a fim de facilitar a visualização das opções de conversão e testes dos resultados por meio do *ebook reader* disponível após a instalação do *software*.

3. Utilização do Calibre

Para utilização do Calibre, tanto na versão visual ou por linha de comando, basta fazer o *download* da aplicação para o sistema operacional da máquina por meio do *site* (https://calibre-ebook.com/download) e seguir os passos de instalação conforme solicitado pelo instalador.

3.1 Instalação do *software*

Após realizada a instalação, basta abrir o *software* Calibre.

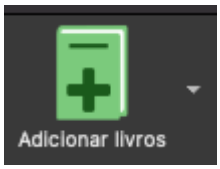


Através da interface visual, as editoras poderão adicionar suas obras e testar as conversões disponíveis.

3.2 Adicionando uma nova obra

Seguindo o padrão determinado para o projeto PNLD Digital, todas as obras deverão seguir o padrão HTML5 e ser dispostos em um único arquivo HTML ou em um conjunto de arquivos que podem ser do tipo HTML para marcação do texto, CSS para folha de estilos, imagens, vídeos e / ou arquivos de configuração.

Para adicionar uma nova obra, basta selecionar a opção de adicionar livros e escolher os arquivos desejados.

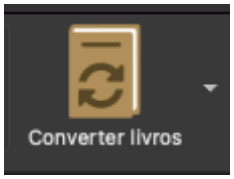


O Calibre permite que se escolha o arquivo "index.html" diretamente, um conjunto de arquivos ou pastas através da interface gráfica ou um arquivo compactado no formato "ZIP". Para garantir a compatibilidade com o sistema PNLD digital, que deverá receber a obra em um único arquivo compactado no mesmo formato, todos os testes serão importados no Calibre para conversão por meio da compactação no formato ZIP, garantindo assim que o resultado gerado seja igual

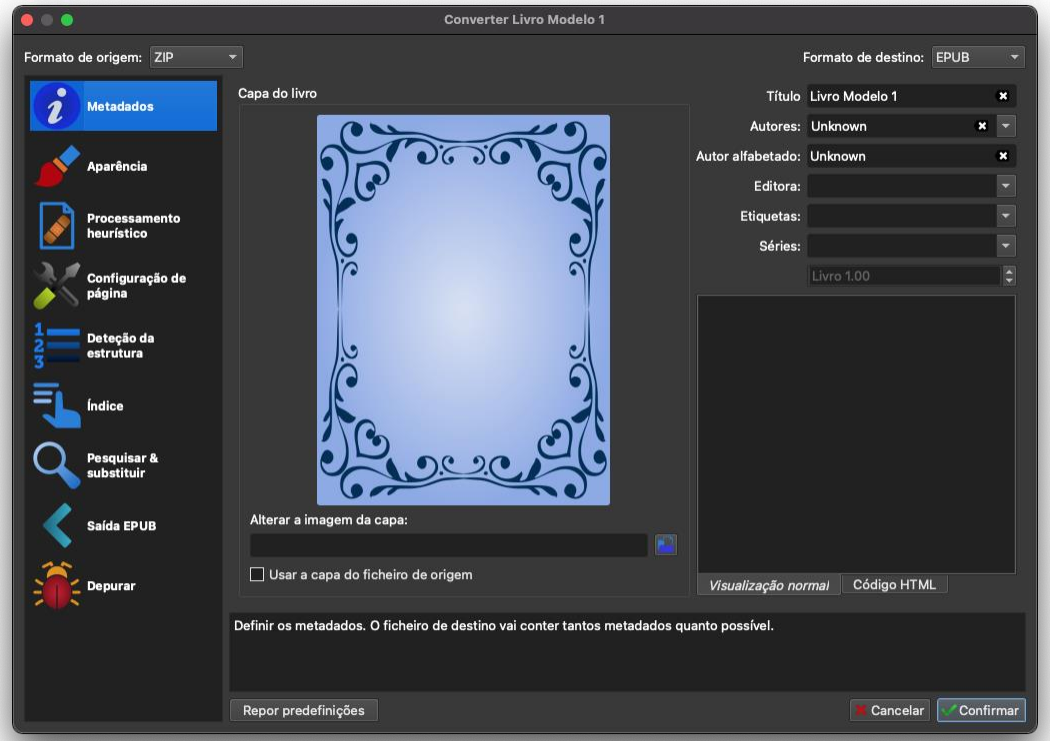
independentemente da plataforma, formato e quantidade de arquivos que compõem o conjunto.

3.3 Conversão das obras.

Para conversão das obras adicionadas à aplicação, basta selecionar a obra desejada e selecionar a opção "Converter livros".

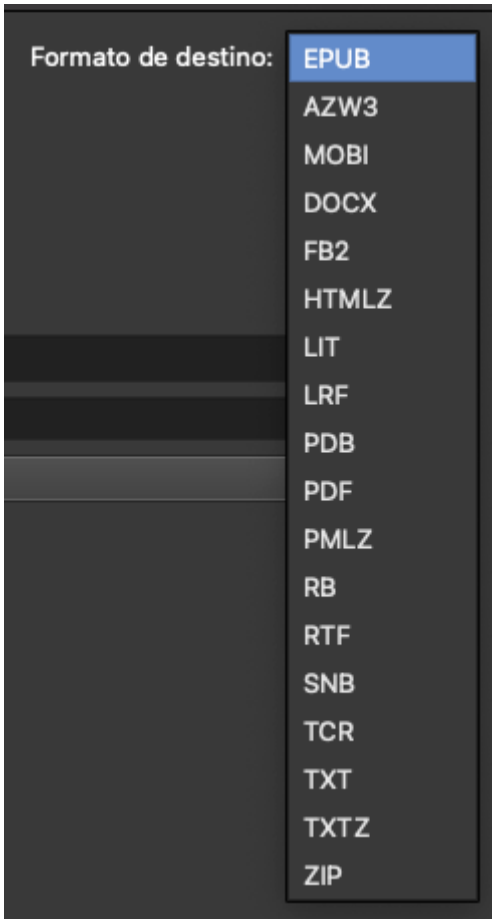


Será aberta uma opção para configuração dos parâmetros de conversão na qual a editora poderá realizar testes específicos da configuração desejada, escolher o formato de saída e editar informações importantes do processo.

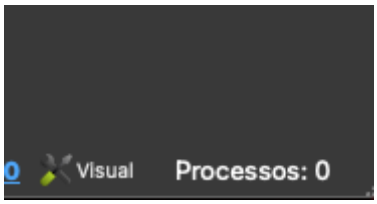


Para efeito da conversão automatizada, que será realizada pelo PNLD Digital automaticamente após validação do sistema para distribuição das obras, as editoras deverão se ater às configurações apenas do formato de destino para escolha do tipo de arquivo ao qual a obra deve ser convertida.

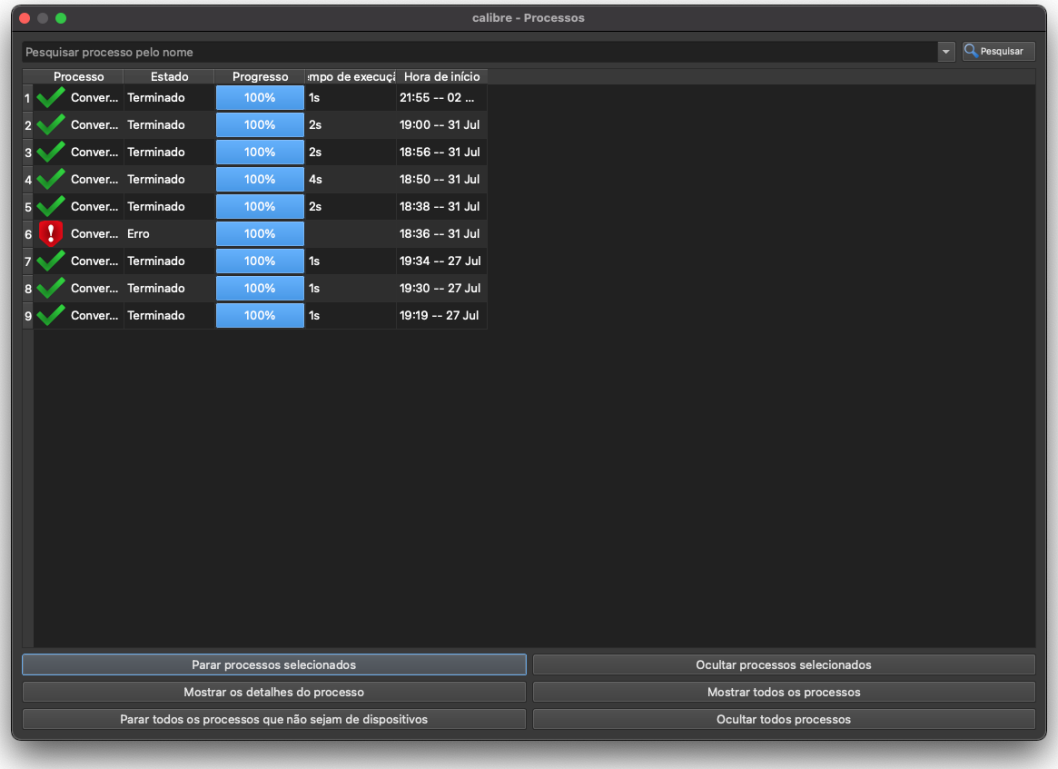
Na versão 5.23.0, utilizada nos testes, é possível fazer a conversão de uma obra desenvolvida no formato HTML5 e submetida por meio de um arquivo compactado no formato "ZIP" para os seguintes formatos: EPUB, AZW3, MOBI, DOCX, FB2, HTMLZ, LIT, LRF, PDB, PDF, PMLZ, RB, RTF, SNB, TCR, TXT e ZIP.



Após confirmar a conversão no formato escolhido, a aplicação iniciará um processo de conversão da obra que poderá ser acompanhada por meio da janela de processos que poderá ser aberta selecionando a identificação de processos no canto inferior direito da aplicação.

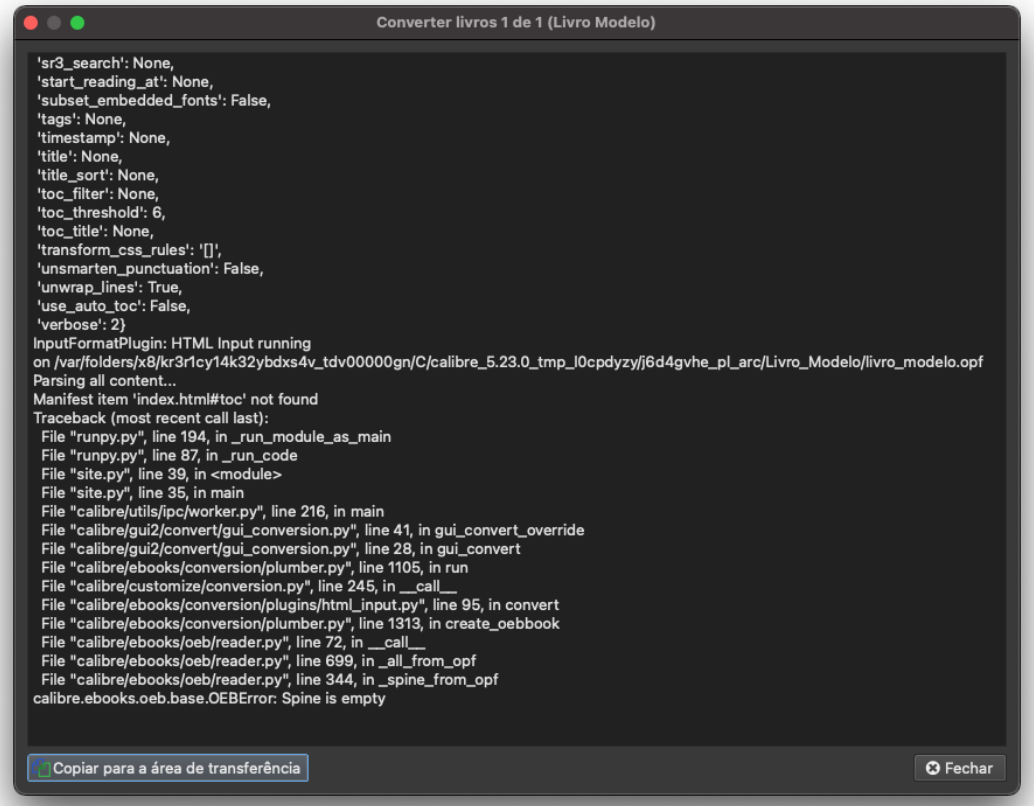


Nessa janela, serão apresentados os processos em andamento, finalizados e se o processo apresentou algum erro na etapa de conversão da obra.

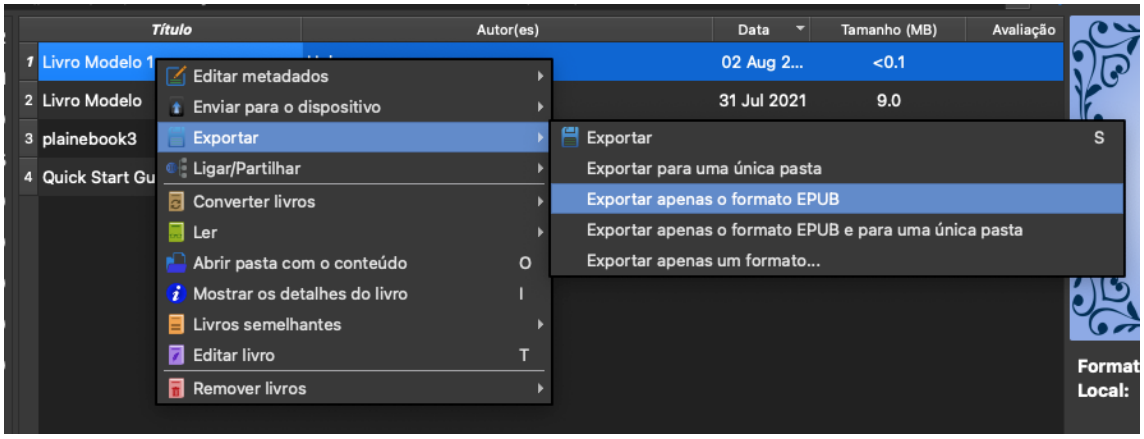


Essa janela é importante para que as editoras possam validar e verificar se houve algum erro no processo de conversão da obra e como resolver os problemas encontrados.

Como exemplo, uma obra que apresentou uma falha na conversão. O erro se caracterizou pela ausência da tag "Spine" obrigatória ao formato, no arquivo OPF, que contém os dados da obra, .



Por fim, para obter a obra convertida, basta exportar no formato desejado.



Nesse momento, o arquivo convertido será salvo na pasta escolhida e poderá ser validado em qualquer leitor digital que implemente o formato escolhido.

3.4 Utilizando a CLI

O Calibre oferece uma série de ferramentas que podem ser utilizadas por meio da linha de comando (Command-line interface - CLI), isso permite a automação e/ou integração desses serviços em *software* de terceiros permitindo, por exemplo, a conversão em lote de arquivos.

Todas as ferramentas estão disponíveis na pasta de instalação do Calibre e estão devidamente documentadas no link https://manual.calibre-ebook.com/generated/en/cli-index.html.

Para os testes de conversão das obras, podemos utilizar tanto a interface gráfica disponibilizada pelo Calibre ou a ferramenta *ebook-converter* que, inclusive, possui mais opções de parametrização de entrada e saída que a versão gráfica como pode ser observado na documentação disponível em https://manual.calibre-ebook.com/generated/en/ebook-convert.html.

4. Formatos utilizados

Para realização dos testes de conversão serão desenvolvidas algumas obras modelo, utilizando as tecnologias HTML e CSS, conforme definido a partir de todo o processo de discussão com o mercado . Como uma obra completa é composta por diversos arquivos e recursos, a obra modelo será compactada no padrão ZIP para realização da conversão.

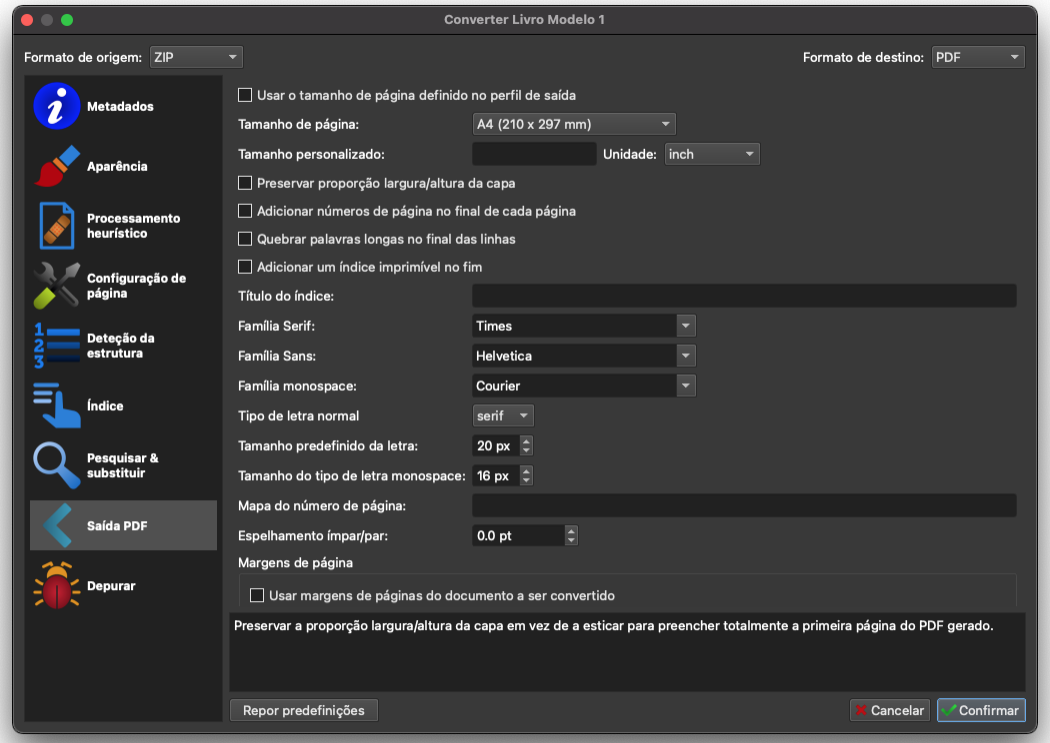
Destaca-se quatro formatos de saída que serão avaliados nos testes, sendo:

- EPUB: Formato mais utilizado atualmente na distribuição de livros digitais, aceito na maioria dos leitores.
- MOBI: Formato proprietário da Amazon para os leitores Kindle, sendo a Amazon uma das maiores distribuidoras de livros digitais da atualidade.
- AZW3: Evolução do padrão MOBI para os dispositivos Kindle, que permite a utilização de recursos interativos e diagramação mais sofisticadas nos dispositivos da empresa.
- PDF: Principal formato utilizado na distribuição de conteúdo digital.

Ao realizar os testes com os formatos especificados, podemos garantir que as obras podem ser convertidas por meio do Calibre para os principais formatos aceitos atualmente pelos leitores digitais, sendo EPUB, MOBI e PDF os principais formatos e AZW3 uma variação do formato disponível para os leitores Kindle (Amazon) com suporte a elementos multimídia disponíveis no HTML5.

4.1 Configuração de saída do formato PDF

Para realizar a conversão para um arquivo PDF, é importante salientar que o formato deve ser configurado previamente, principalmente com relação ao tamanho da página e suas margens, já que esa configuração pode influenciar no resultado final.



4.2 Validação de saída do formato EPUB

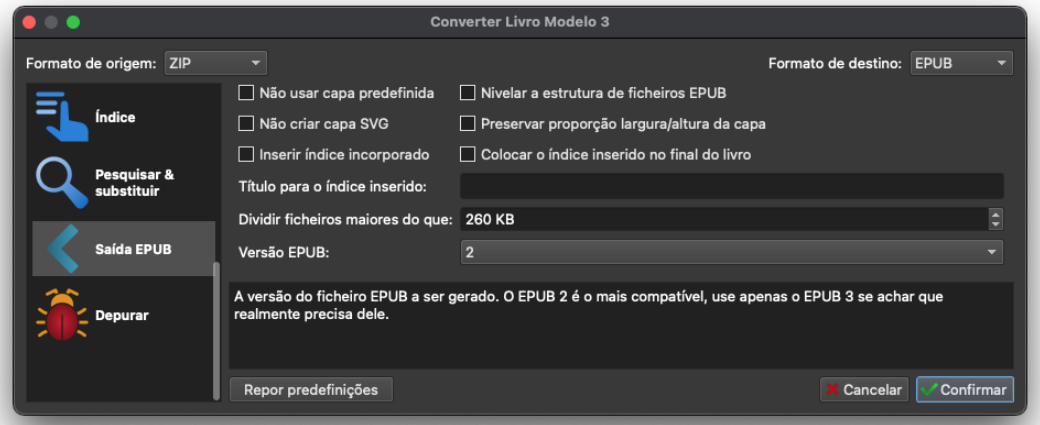
O formato EPUB deve seguir todas as diretrizes convencionadas pela International Digital Publishing Forum - IDPF, mantenedora do formato EPUB em todas as suas versões. Essa validação é importante para garantir que a conversão foi efetuada com sucesso, seguindo todos os padrões exigidos, para que as editoras tenham a liberdade de publicação de sua obra em qualquer publicadora, nacional ou internacional que normalmente exigem esse processo.

Para essa validação da obra, o IDPF disponibiliza uma ferramenta *online* em http://validator.idpf.org/, na qual a obra é analisada. Para obras que possuem dados sensíveis ou sigilo de informação, é possível fazer a instalação da ferramenta de validação em um equipamento local, evitando assim submeter a seu conteúdo a um *site* na internet.

Importante ressaltar, também, que o EPUB possui duas versões de saída comumente vista no mercado sendo: EPUB 2 e EPUB 3. Ambas as versões apresentam resultados satisfatórios para a

leitura e apresentação de conteúdo digital. Porém, a versão EPUB 3 possui a compatibilidade com HTML5 e CSS3, o que garante algumas vantagens ao formato, como a utilização das tags semânticas provenientes do HTML5, por exemplo, *<article>*, *<section>*, *<nav>* entre outras, o que possibilita uma melhor implementação quanto à acessibilidade da obra.

Em contrapartida, conforme descrito pelo Calibre, a versão EPUB 2 é uma forma que garante uma melhor compatibilidade entre os diversos equipamentos e leitores disponíveis no mercado, mas que demanda a escrita ser feita utilizando de tecnologias mais antigas.



Na realização dos testes, vamos buscar inicialmente a maior interoperabilidade entre os formatos disponíveis e, por esse motivo, o resultado final será comparado inicialmente com o formato EPUB2. Porém, os testes voltados à parte de acessibilidade deverão ser mais profundos com relação aos dois formatos a fim de identificar a diferença entre eles na utilização do HTML5.

5. Realização dos testes

Para realização dos testes, foram criadas algumas obras com um conjunto de informações para cada validação e/ou tipo de teste.

Todas as obras foram criadas tendo como base o padrão de HTML definidos a partir de estudos e da coleta de informações do mercado, com as devidas tags e metadados para identificação da obra e seu conteúdo.

5.1 Livro Modelo 1 - Obra simples em arquivo único HTML.

Teste voltado à conversão das tags de marcação de texto presentes no HTML para validar os textos da obra e como serão apresentados após a conversão.

5.1.1 Preparação da obra

Foi desenvolvida uma obra de arquivo único, "index.html", denominado "Livro_Modelo_1".

Esse modelo apresenta no corpo do documento uma área de navegação com a Tabela de Conteúdo (TOC) e uma área principal com elementos textuais e tags HTML com grande representatividade na elaboração de obra literária como: *h1, h2, h3, p, table*.

Todo o conteúdo foi disposto em três capítulos. Em cada capítulo há três seções internas para navegação, sendo que, ao final do capítulo 1 e 3, há uma tabela simples para validação da apresentação dos dados.

Foi adicionado uma tabela com largura de 100% da dimensão de destino para validar o comportamento na renderização de um componente mais complexo, que deve se adaptar às múltiplas telas possíveis em dispositivos ou até mesmo ao tamanho de folhas no formato PDF.

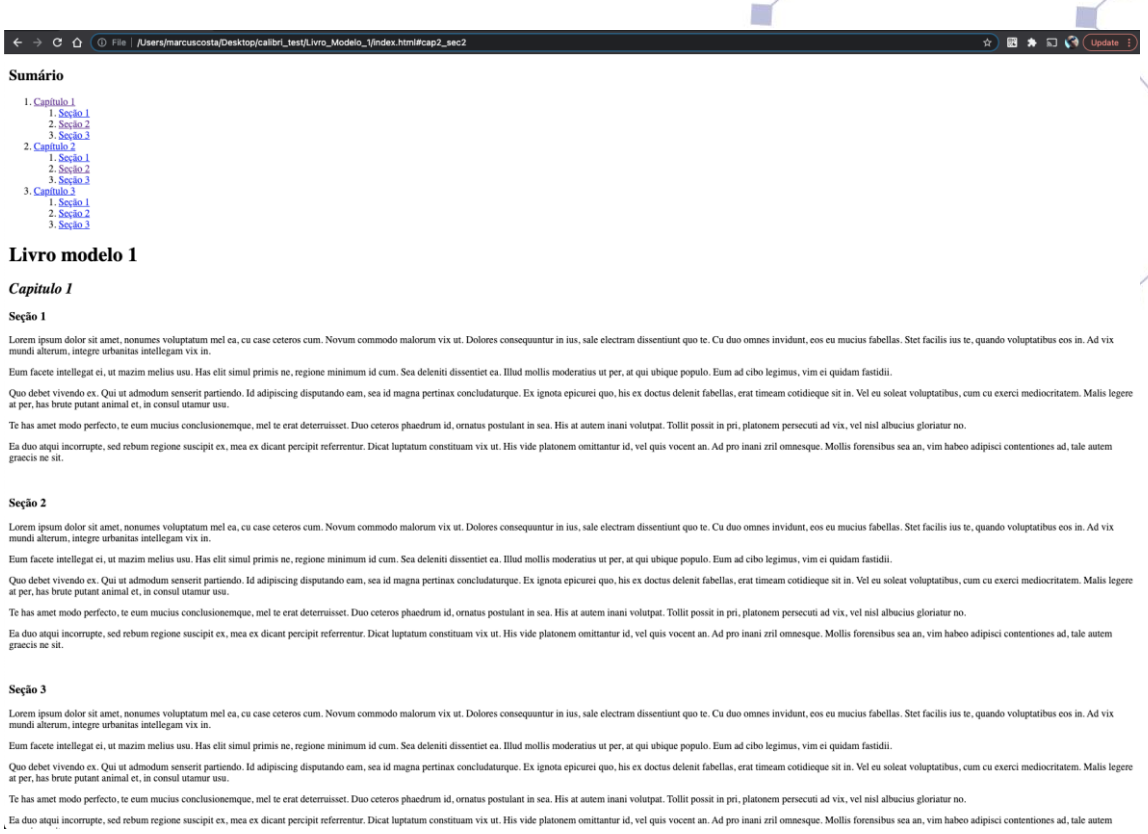
5.1.2 Objetivo

Observar o comportamento dos textos nos leitores digitais, validando a troca de fonte, acompanhamento do conteúdo (porcentagem lida da obra) e a navegação entre os diferentes capítulos e seções.

Validação da renderização do conteúdo em múltiplos formatos e telas de tamanhos diferentes.

5.1.3 Apresentação da obra original - HTML

A obra original se apresenta de uma maneira bem direta em qualquer navegador de internet, com uma sessão de sumário capaz de navegar normalmente entre os capítulos e seções disponíveis.



File | /Users/marcuscosta/Desktop/calibri_test/Livro_Modelo_1/index.html#cap2_sec2

Update

Sumário

1. Capítulo 1
    1. Seção 1
    2. Seção 2
    3. Seção 3
2. Capítulo 2
    1. Seção 1
    2. Seção 2
    3. Seção 3
3. Capítulo 3
    1. Seção 1
    2. Seção 2
    3. Seção 3

Livro modelo 1

Capitulo 1

Seção 1

Lorem ipsum dolor sit amet, nonumes voluptatum mel ea, cu case ceteros cum. Novum commodo malorum vix ut. Dolores consequuntur in ius, sale electram dissentiunt quo te. Cu duo omnes invidunt, eos eu mucius fabellas. Stet facilis ius te, quando voluptatibus eos in. Ad vix mundi alterum, integre urbanitas intellegam vix in.

Eum facete intellegat ei, ut mazim melius usu. Has elit simul primis ne, regione minimum id cum. Sea deleniti dissentiet ea. Illud mollis moderatius ut per, at qui ubique populo. Eum ad cibo legimus, vim ei quidam fastidii.

Quo debet vivendo ex. Qui ut admodum senserit partiendo. Id adipiscing disputando eam, sea id magna pertinax concludaturque. Ex ignota epicurei quo, his ex doctus delenit fabellas, erat timeam cotidieque sit in. Vel eu soleat voluptatibus, cum cu exerci mediocritatem. Malis legere at per, has brute putant animal et, in consul utamur usu.

Te has amet modo perfecto, te eum mucius conclusionemque, mel te erat deterruisset. Duo ceteros phaedrum id, ornatus postulant in sea. His at autem inani volutpat. Tollit possit in pri, platonem persecuti ad vix, vel nisl albucius gloriatur no.

Ea duo atqui incorrupte, sed rebum regione suscipit ex, mea ex dicant percipit referrentur. Dicat luptatum constituam vix ut. His vide platonem omittantur id, vel quis vocent an. Ad pro inani zril omnesque. Mollis forensibus sea an, vim habeo adipisci contentiones ad, tale autem graecis ne sit.

Seção 2

Lorem ipsum dolor sit amet, nonumes voluptatum mel ea, cu case ceteros cum. Novum commodo malorum vix ut. Dolores consequuntur in ius, sale electram dissentiunt quo te. Cu duo omnes invidunt, eos eu mucius fabellas. Stet facilis ius te, quando voluptatibus eos in. Ad vix mundi alterum, integre urbanitas intellegam vix in.

Eum facete intellegat ei, ut mazim melius usu. Has elit simul primis ne, regione minimum id cum. Sea deleniti dissentiet ea. Illud mollis moderatius ut per, at qui ubique populo. Eum ad cibo legimus, vim ei quidam fastidii.

Quo debet vivendo ex. Qui ut admodum senserit partiendo. Id adipiscing disputando eam, sea id magna pertinax concludaturque. Ex ignota epicurei quo, his ex doctus delenit fabellas, erat timeam cotidieque sit in. Vel eu soleat voluptatibus, cum cu exerci mediocritatem. Malis legere at per, has brute putant animal et, in consul utamur usu.

Te has amet modo perfecto, te eum mucius conclusionemque, mel te erat deterruisset. Duo ceteros phaedrum id, ornatus postulant in sea. His at autem inani volutpat. Tollit possit in pri, platonem persecuti ad vix, vel nisl albucius gloriatur no.

Ea duo atqui incorrupte, sed rebum regione suscipit ex, mea ex dicant percipit referrentur. Dicat luptatum constituam vix ut. His vide platonem omittantur id, vel quis vocent an. Ad pro inani zril omnesque. Mollis forensibus sea an, vim habeo adipisci contentiones ad, tale autem graecis ne sit.

Seção 3

Lorem ipsum dolor sit amet, nonumes voluptatum mel ea, cu case ceteros cum. Novum commodo malorum vix ut. Dolores consequuntur in ius, sale electram dissentiunt quo te. Cu duo omnes invidunt, eos eu mucius fabellas. Stet facilis ius te, quando voluptatibus eos in. Ad vix mundi alterum, integre urbanitas intellegam vix in.

Eum facete intellegat ei, ut mazim melius usu. Has elit simul primis ne, regione minimum id cum. Sea deleniti dissentiet ea. Illud mollis moderatius ut per, at qui ubique populo. Eum ad cibo legimus, vim ei quidam fastidii.

Quo debet vivendo ex. Qui ut admodum senserit partiendo. Id adipiscing disputando eam, sea id magna pertinax concludaturque. Ex ignota epicurei quo, his ex doctus delenit fabellas, erat timeam cotidieque sit in. Vel eu soleat voluptatibus, cum cu exerci mediocritatem. Malis legere at per, has brute putant animal et, in consul utamur usu.

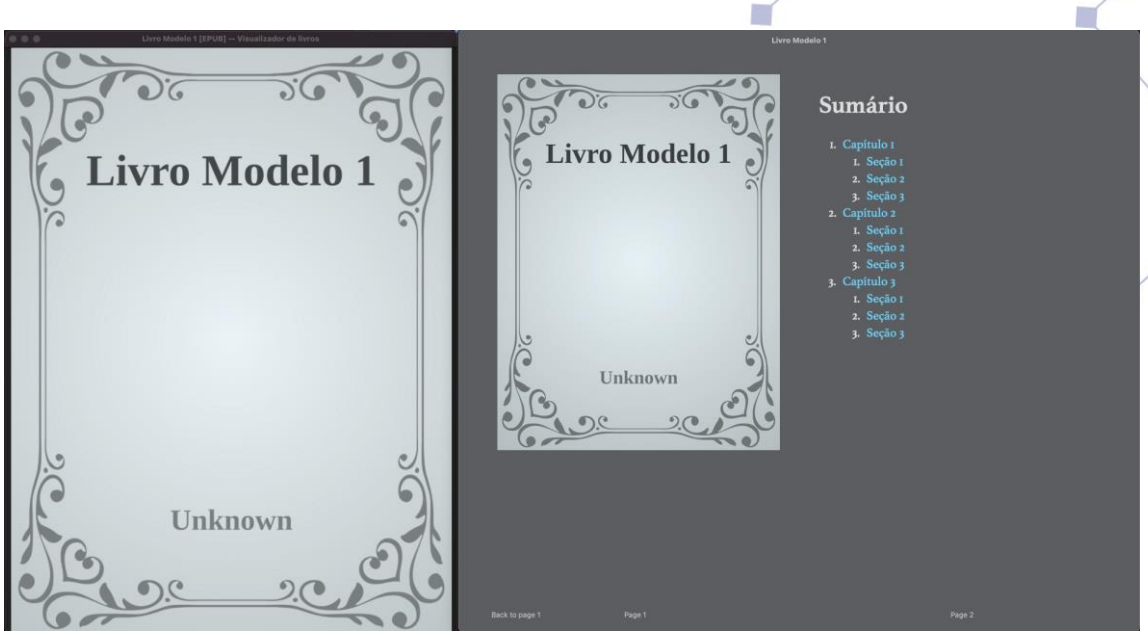
Te has amet modo perfecto, te eum mucius conclusionemque, mel te erat deterruisset. Duo ceteros phaedrum id, ornatus postulant in sea. His at autem inani volutpat. Tollit possit in pri, platonem persecuti ad vix, vel nisl albucius gloriatur no.

Ea duo atqui incorrupte, sed rebum regione suscipit ex, mea ex dicant percipit referrentur. Dicat luptatum constituam vix ut. His vide platonem omittantur id, vel quis vocent an. Ad pro inani zril omnesque. Mollis forensibus sea an, vim habeo adipisci contentiones ad, tale autem

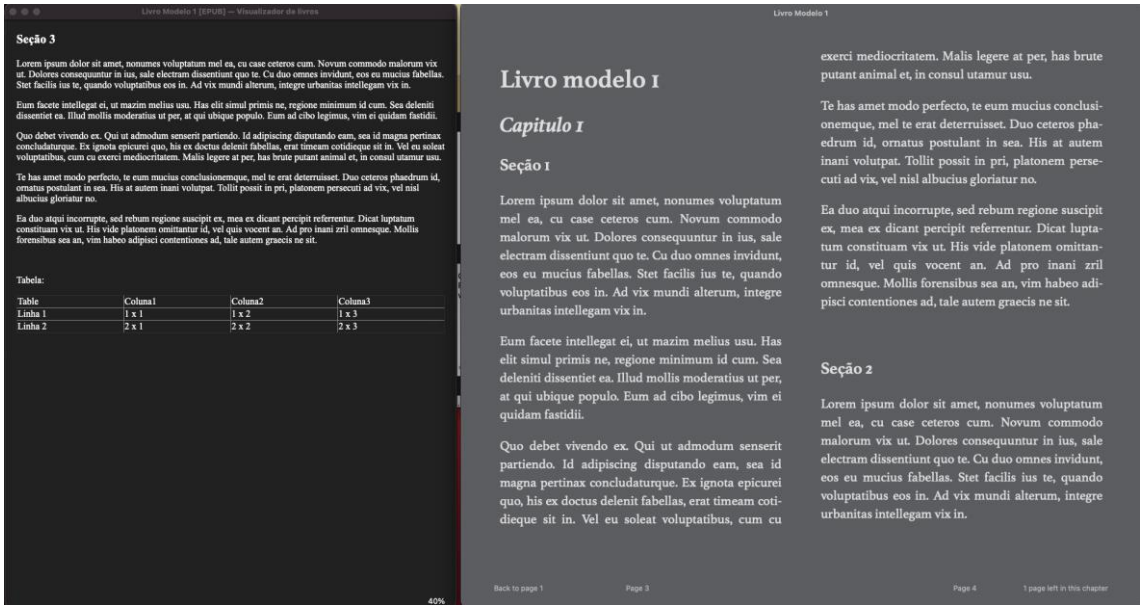
5.1.4 Conversão para EPUB

O formato apresentou um bom resultado de saída, tanto no Ebook-Viewer quanto no iBooks, sendo que o conversor adiciona automaticamente uma capa padrão na obra, que pode ser configurada.

Ambos os leitores apresentaram um sumário navegável entre as âncoras das seções e capítulos.



Todas as funcionalidades dispostas em cada leitor funcionaram perfeitamente como a troca de fonte, tamanho do texto, porcentagem de leitura (Ebook-Viewer), contagem de páginas para fim do capítulo (iBooks), sendo que o conversor adicionou automaticamente a quebra de páginas antes da tag h1, principal elemento do HTML para identificar o *Head* principal.



A tabela, que deve seguir a porcentagem do tamanho da tela, também foi renderizada da forma correta, mesmo efetuando a alteração de fonte e de seu tamanho.

Um ponto de falha importante a ser discutido é que a obra convertida não foi validada pelo **IDPF Validator**. Apresentando falhas nas tags HTML que não são suportadas na convenção do IPDF para o formato.

Not Secure | validator.idpf.org/application/validate

EPUB Validator (beta)

Results

Detected version: EPUB 2.0.1

Results: The following problems were found in *Livro Modelo 1 - Unknown.epub*:

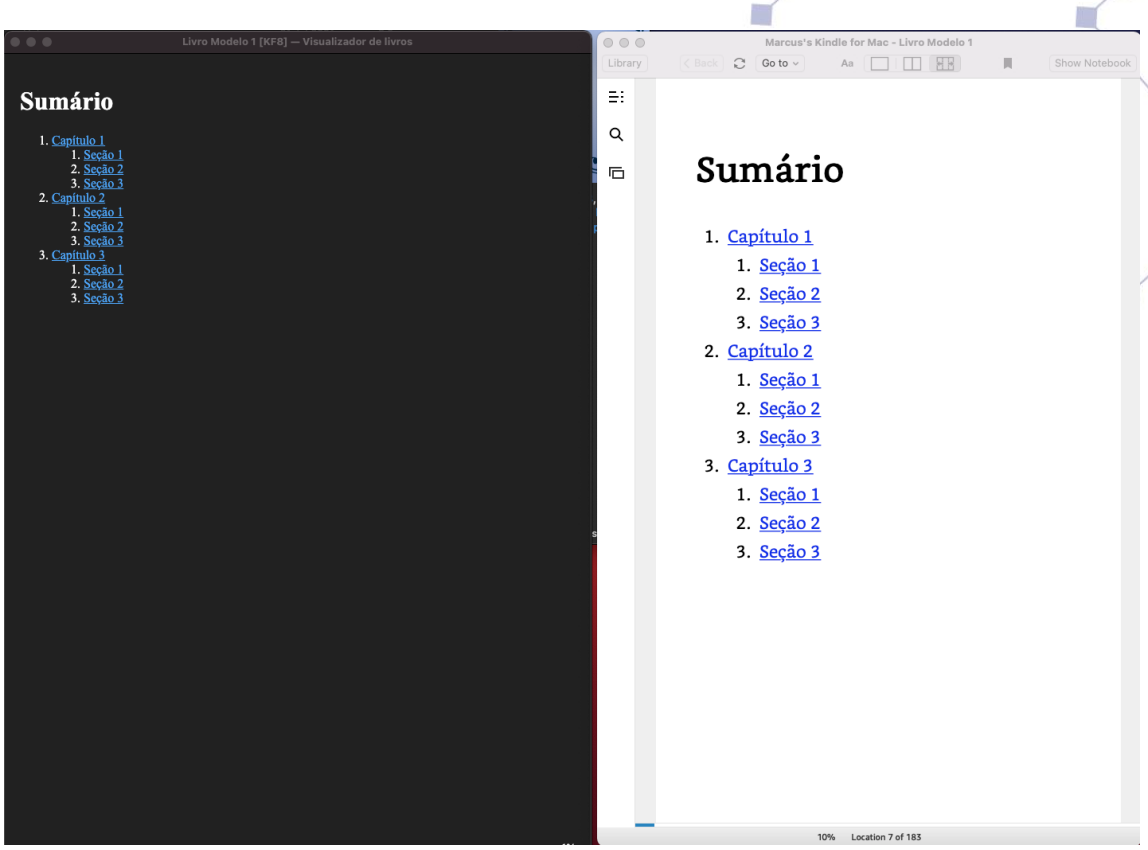
Type	File	Line	Position	Message
ERROR	index_split_000.html	14	30	Error while parsing file: element "nav" not allowed anywhere; expected element "address", "blockquote", "del", "div", "dl", "h1", "h2", "h3", "h4", "h5", "h6", "hr", "ins", "noscript", "ns:svg", "ol", "p", "pre", "script", "table" or "ul" (with xmlns:ns="http://www.w3.org/2000/svg")
ERROR	index_split_000.html	15	58	Error while parsing file: element "nav" not allowed anywhere; expected element "address", "blockquote", "del", "div", "dl", "h1", "h2", "h3", "h4", "h5", "h6", "hr", "ins", "noscript", "ns:svg", "ol", "p", "pre", "script", "table" or "ul" (with xmlns:ns="http://www.w3.org/2000/svg")
ERROR	index_split_000.html	41	12	Error while parsing file: element "nav" not allowed anywhere; expected element "address", "blockquote", "del", "div", "dl", "h1", "h2", "h3", "h4", "h5", "h6", "hr", "ins", "noscript", "ns:svg", "ol", "p", "pre", "script", "table" or "ul" (with xmlns:ns="http://www.w3.org/2000/svg")
ERROR	index_split_000.html	43	12	Error while parsing file: element "body" incomplete; expected element "address", "blockquote", "del", "div", "dl", "h1", "h2", "h3", "h4", "h5", "h6", "hr", "ins", "noscript", "ns:svg", "ol", "p", "pre", "script", "table" or "ul" (with xmlns:ns="http://www.w3.org/2000/svg")
ERROR	index_split_001.html	14	7	Error while parsing file: element "main" not allowed anywhere; expected element "address", "blockquote", "del", "div", "dl", "h1", "h2", "h3", "h4", "h5", "h6", "hr", "ins", "noscript", "ns:svg", "ol", "p", "pre", "script", "table" or "ul" (with xmlns:ns="http://www.w3.org/2000/svg")
ERROR	index_split_001.html	15	99	Error while parsing file: element "main" not allowed anywhere; expected element "address", "blockquote", "del", "div", "dl", "h1", "h2", "h3", "h4", "h5", "h6", "hr", "ins", "noscript", "ns:svg", "ol", "p", "pre", "script", "table" or "ul" (with xmlns:ns="http://www.w3.org/2000/svg")
				Error while parsing file: element "main" not allowed anywhere; expected element "address", "blockquote", "del", "div", "dl", "h1", "h2"

As falhas apresentadas foram referentes às tags de navegação *<nav>* e *<main>* dispostas no arquivo HTML para navegação no conteúdo e criação do *TOC* que não fazem parte do padrão EPUB na versão 2, sendo que para a versão 3 a validação ocorre normalmente.

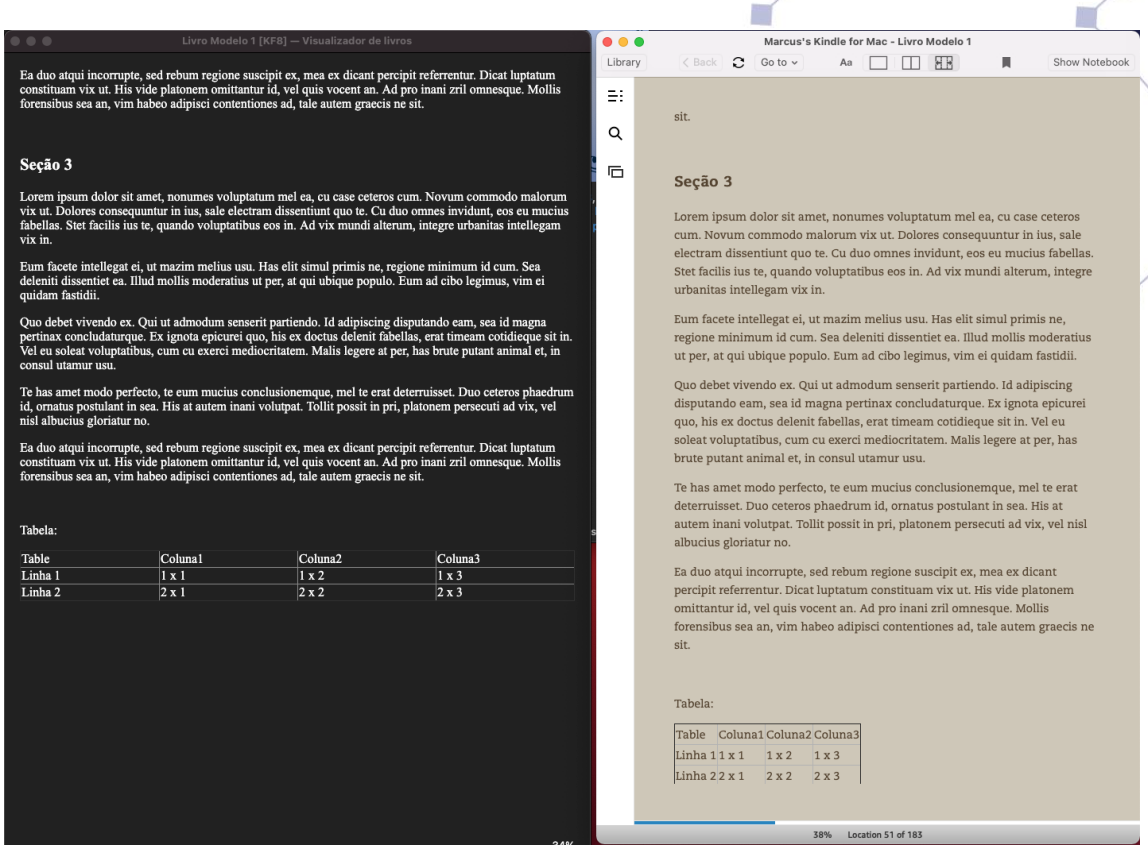
Esse modelo foi construído a partir do compilado de estudos e das contribuições dos diversos produtores do mercado que têm enviado documentos e sugestões para a criação do PNLD Digital. Com base nesses estudos, erá necessária a alteração da disposição desses elementos para criação de um arquivo *NXC* para a criação da tabela de conteúdo e OPF para mapeamento das informações da obra. Os testes referentes a essa abordagem serão discutidos no item 5.2 Livro Modelo 2 - Obra com dados e capa (Arquivo OPF).

5.1.5 Conversão para MOBI e AZW3

O formato apresentou um bom resultado de saída, tanto no Ebook-Viewer quanto no Kindle e ambos os leitores apresentaram um sumário navegável entre as âncoras das seções e capítulos.



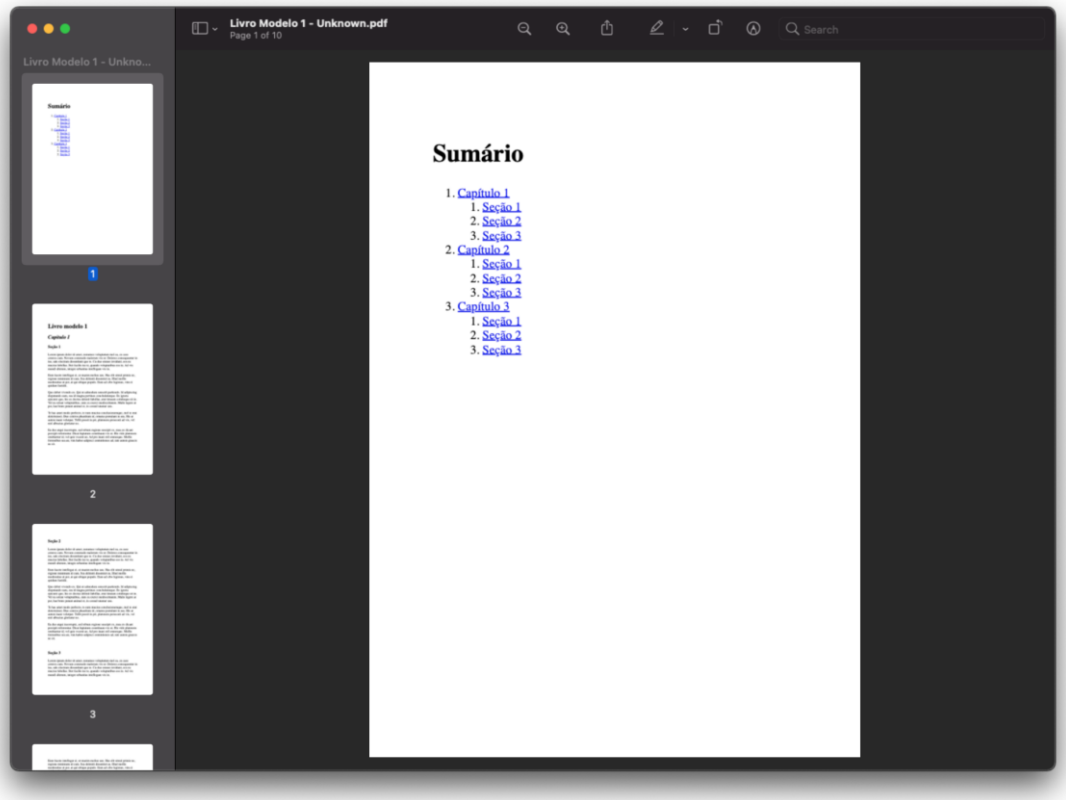
Todas as funcionalidades dispostas em cada leitor funcionaram perfeitamente como a troca de fonte, tamanho do texto, porcentagem de leitura, sendo que o conversor adicionou automaticamente a quebra de páginas antes da tag h1, principal elemento do HTML para identificar o *Head* principal.



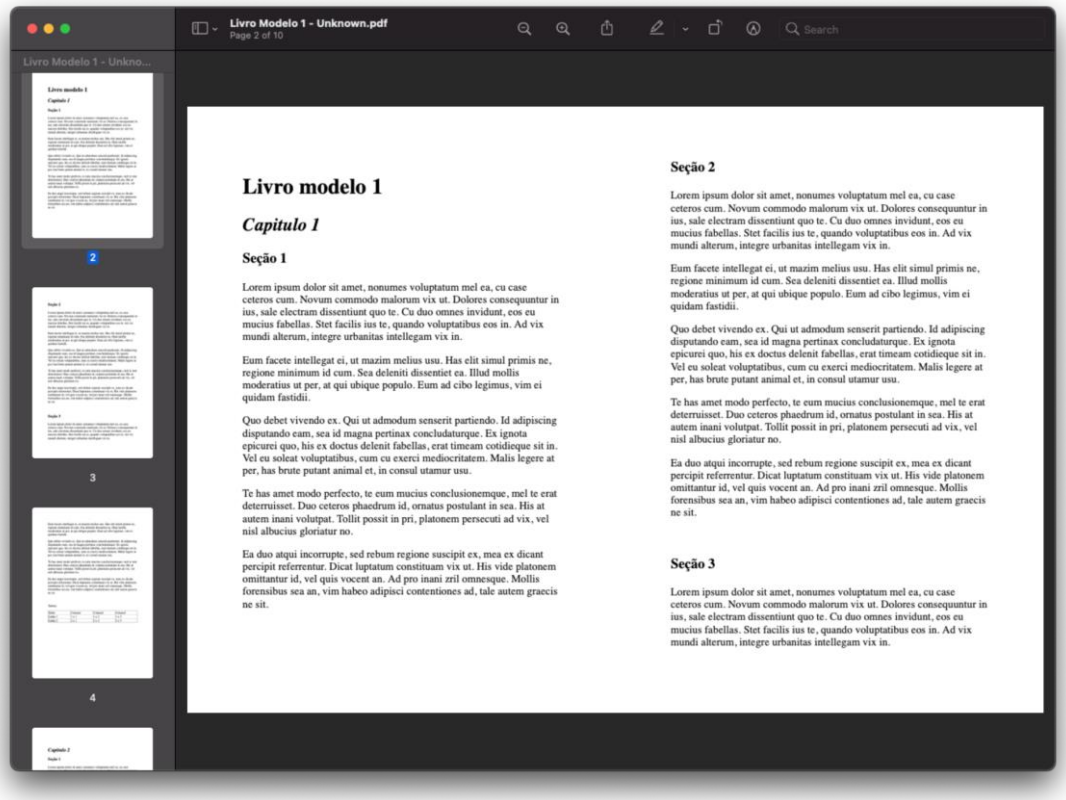
Porém o Kindle não renderiza a tabela com o tamanho correto, seguindo os parâmetros dispostos no HTML para apresentação e utilização de 100% da largura disponível da tela.

5.1.6 Conversão para PDF

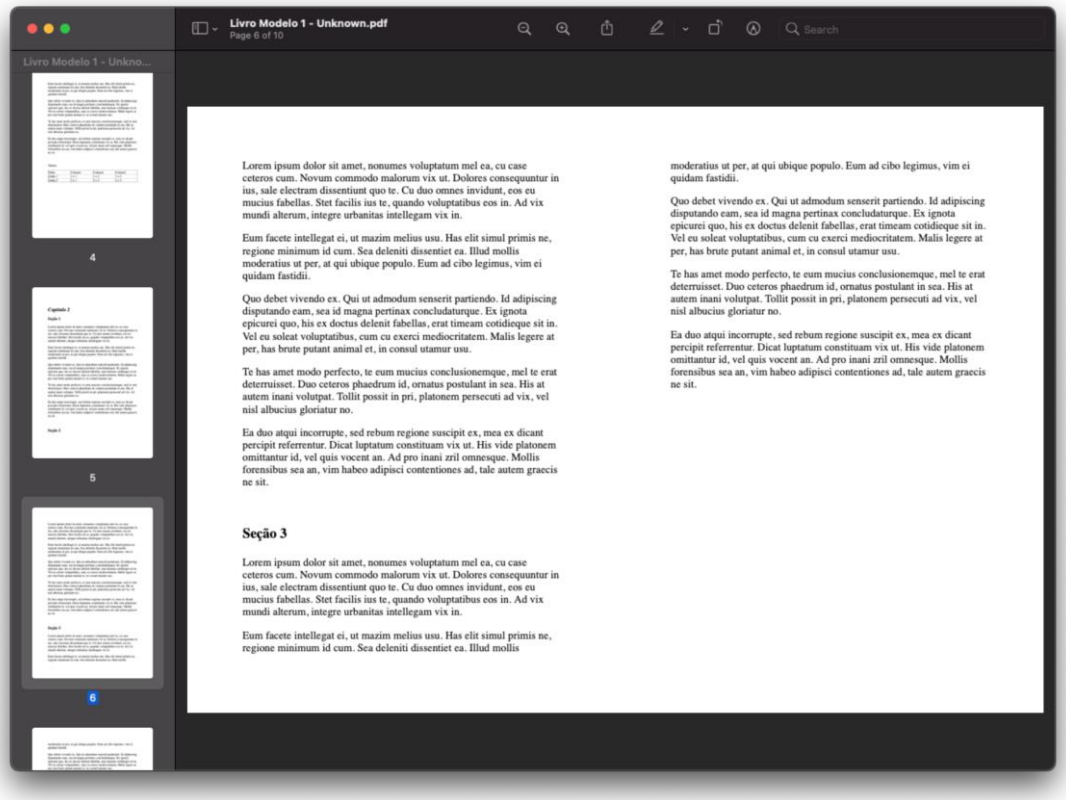
O Formato apresentou um bom resultado de saída, sendo a página inicial com o sumário navegável entre as âncoras das seções e capítulos.



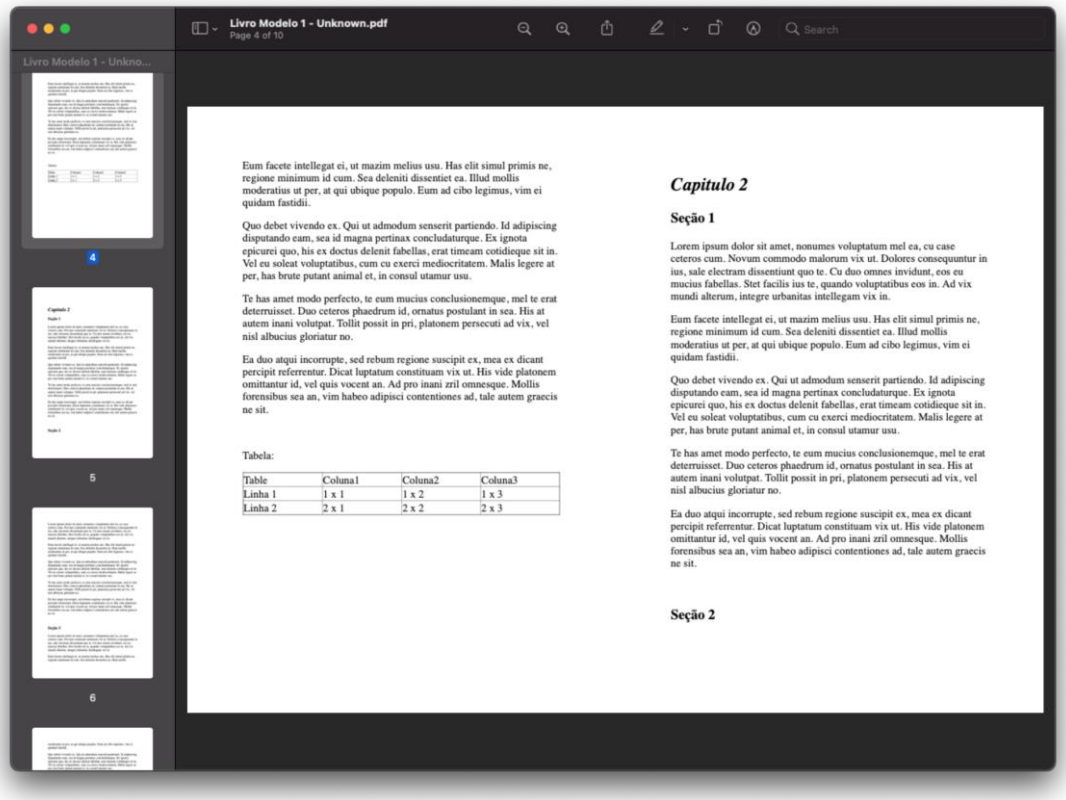
Os textos ficaram legíveis e devidamente dispostos nas páginas de acordo com a configuração de tamanho da página e margem escolhidos.



Antes de cada tag h1, principal elemento de indicação de um novo *header* na obra, o conversor adiciona automaticamente uma quebra de página, garantindo uma melhor organização do conteúdo entre os capítulos.



Elementos complexos, como tabelas, foram renderizadas corretamente e utilizaram todo o espaço disponível na tela.

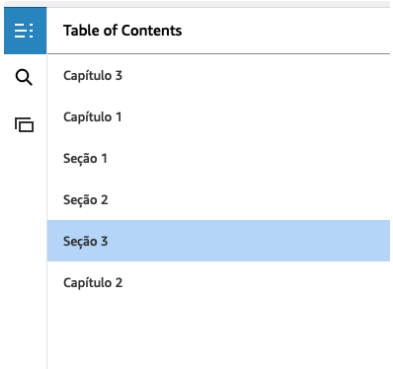


5.1.7 Conclusão: Livro Modelo 1

Observando os resultados obtidos em cada um dos leitores, o processo de conversão apresenta um resultado satisfatório para a apresentação de textos e navegação do conteúdo.

Porém, pode-se observar que a renderização final do conteúdo poderá ser diferente em cada tipo de leitor, devendo a editora estar ciente desse processo.

Único problema encontrado nos testes, para esse exemplo inicial, encontra-se na tabela de conteúdo, gerada automaticamente pelo conversor que é utilizado na navegação. Apesar de apresentar o sumário navegável corretamente, todos os leitores apresentam uma forma de navegação através do elemento ToC (Table of Contents) e em todos os casos não apresentou todas as seções de forma adequada e, no caso do Kindle, ficou com a ordem do conteúdo errada.



5.2 Livro Modelo 2 - Obra com metadados e capa (Arquivo OPF)

Teste voltado à conversão das obras com os dados a ela relacionados, como autor, editora, etc, dados da tabela de conteúdo (TOC) e capa do livro.

5.2.1 Preparação da obra

Para a preparação da obra "Livro Modelo 2" foi excluído as tags *<nav>* e *<main>* utilizadas para criação da Tabela de Conteúdo (TOC) devido ao problema observado na validação do formato EPUB, na versão 2, por meio do **IDPF Validator**.

Nesse caso, foi utilizado uma tag como container para separação do sumário e o conteúdo do livro e gerado o arquivo "*toc.nxc*" para descrever a navegação da obra.

Para garantir que todos os metadados fossem importados corretamente da obra original para a conversão, foi implementado na obra o arquivo "content.opf" que descreve todo o conteúdo pertencente à obra, a sua capa e aos metadados.

A construção do arquivo OPF - Open Packaging Format seguiu todas as definições disponibilizadas pelo IDPF no link http://idpf.org/epub/20/spec/OPF_2.0.1_draft.htm, incluindo a configuração da imagem de capa na estrutura da obra e no manifest do arquivo OPF.

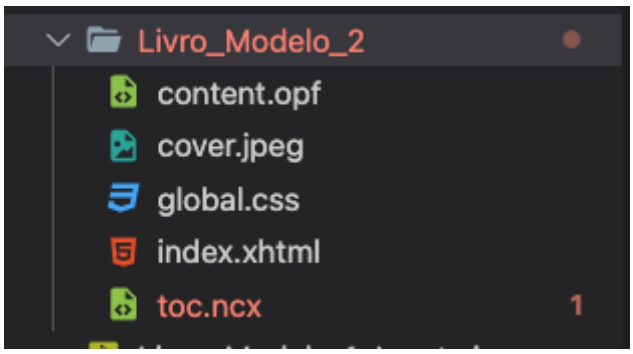
5.2.2 Objetivo

O Calibre se mostrou ineficaz na interpretação dos metadados contidos no HTML ao encapsular a obra no formato ZIP na conversão entre os formatos. Esse teste tem o intuito de validar se o Calibre é capaz de interpretar os metadados e a imagem de capa disponíveis no arquivo modelo original, por meio do arquivo OPF, e inseri-los corretamente no formato convertido.

A troca no modelo de navegação visa a validar se os erros apresentados na geração da tabela de conteúdo (TOC) no teste anterior serão resolvidos nessa versão da obra.

5.2.3 Apresentação da obra original - HTML

A obra original possui o mesmo *layout* do apresentado no "Livro Modelo 1", porém possui mais arquivos em sua estrutura base para ser convertido e interpretado conforme descrito no item 5.2.1.



5.2.4 Validação do formato EPUB

Após as mudanças nos elementos de navegação, a obra convertida para o formato EPUB 2 foi submetida ao IDPF Validator de forma satisfatória.

EPUB Validator (beta)

Results

Detected version: EPUB 2.0.1

Results: Congratulations! No problems were found in *Livro Modelo 2 - Autor PNLD.epub*.

Validated using EPUBCheck version 4.1.1.

Back to validator

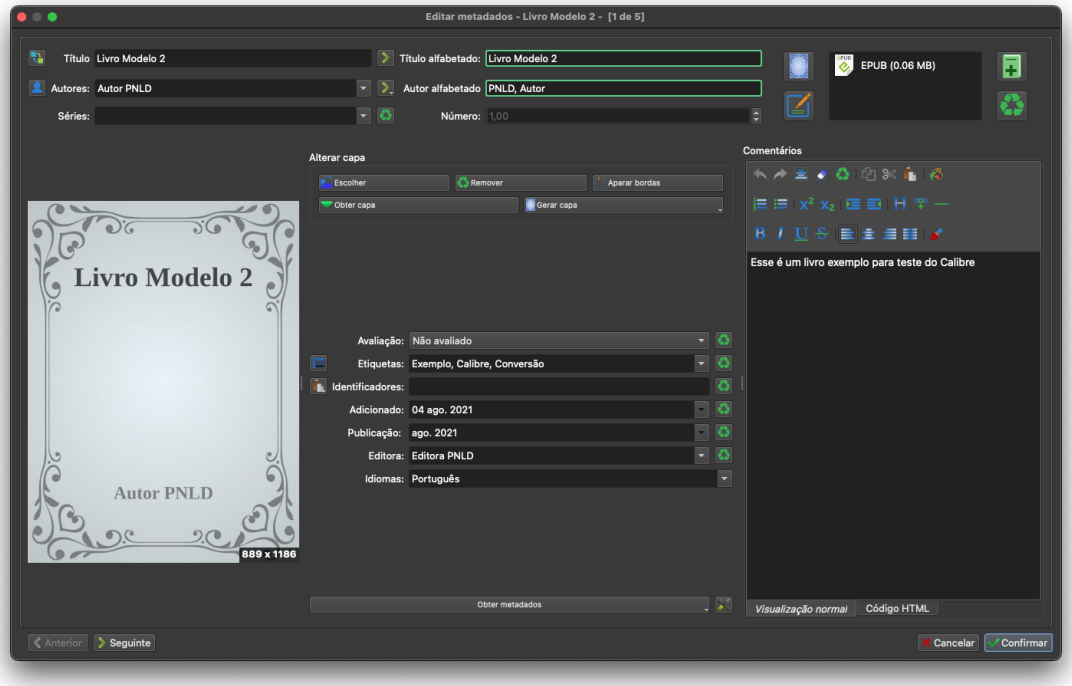
Foi observado que o Calibre não faz a validação das obras e seus formatos, limitando-se a executar a conversão dos dados presentes na obra original para o formato de destino. Nesse caso, cabe ao desenvolvedor da obra garantir que a implementação seja realizada de acordo com a necessidade de cada formato.

Apesar de a obra anterior não ter passado pelo validador do IDPF, não foi identificado nenhuma alteração com relação a apresentação da obra nos leitores testados, em ambas as versões do

EPUB, pois a apresentação depende exclusivamente da implementação do leitor em questão. Algumas publicadoras exigem, contudo, que a obra esteja em total conformidade com as diretrizes de cada formato para que a obra seja disponibilizada.

5.2.5 Observação dos metadados

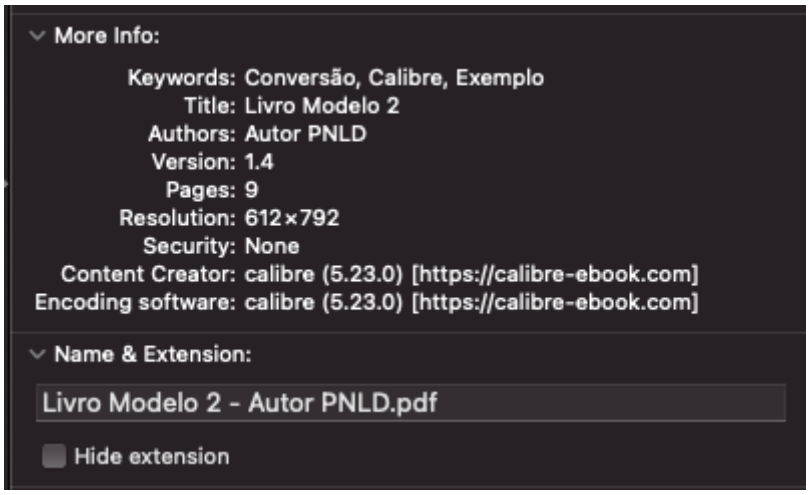
Os metadados adicionados ao arquivo OPF da obra original foram devidamente incorporados ao pacote convertido no formato EPUB, MOBI e AZW3 e podem ser lidos por meio do próprio Calibre, ao adicionar a obra convertida na aplicação.



Na obra convertida para o formato EPUB, podemos analisar o resultado descompactando os dados contidos no pacote EPUB e observando o arquivo OPF final da obra.

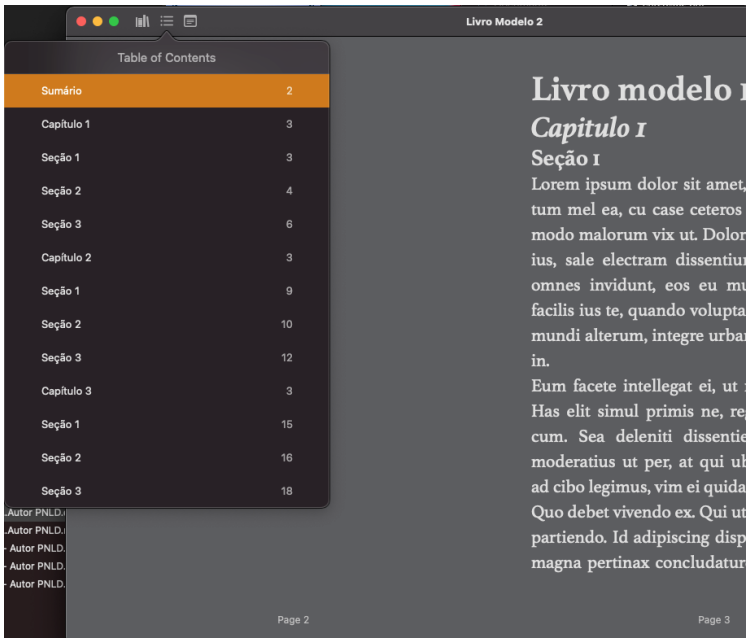
```
<metadata xmlns:opf="http://www.idpf.org/2007/opf" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterm
  <dc:title>Livro Modelo 2</dc:title>
  <meta name="calibre:timestamp" content="2021-08-04T03:01:18.680674+00:00"/>
  <dc:language>pt</dc:language>
  <dc:creator opf:file-as="PNLD, Autor" opf:role="aut">Autor PNLD</dc:creator>
  <dc:identifier id="uuid_id" opf:scheme="uuid">e4e368e6-5d92-4ccf-aa3f-3ef462728abd</dc:identifier>
  <dc:description>Esse é um livro exemplo para teste do Calibre</dc:description>
  <dc:publisher>Editora PNLD</dc:publisher>
  <dc:subject>Exemplo</dc:subject>
  <dc:subject>Calibre</dc:subject>
  <dc:subject>Conversão</dc:subject>
  <dc:date>2021-08-03T03:00:00+00:00</dc:date>
  <dc:rights>Todos os direitos reservados</dc:rights>
  <meta name="cover" content="cover"/>
</metadata>
```

A obra convertida para o formato PDF também apresentou resultados satisfatórios ao manter corretamente os dados inseridos na obra original ao arquivo PDF.

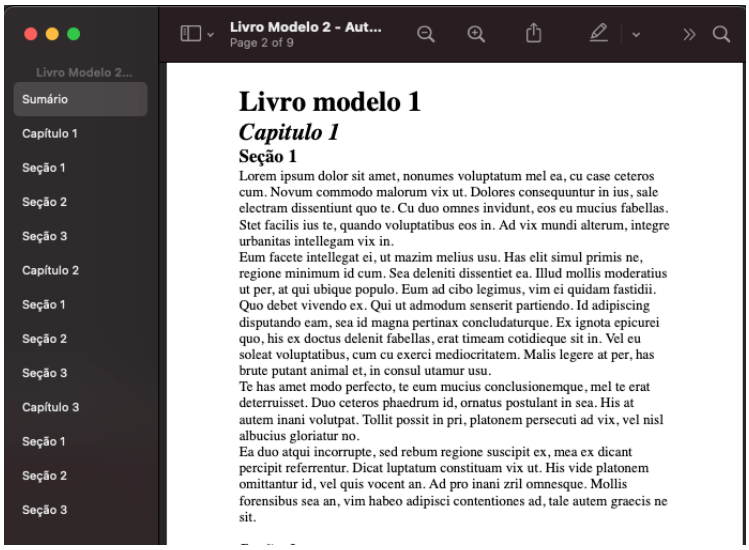


5.2.6 Observação da Tabela de Conteúdo (TOC)

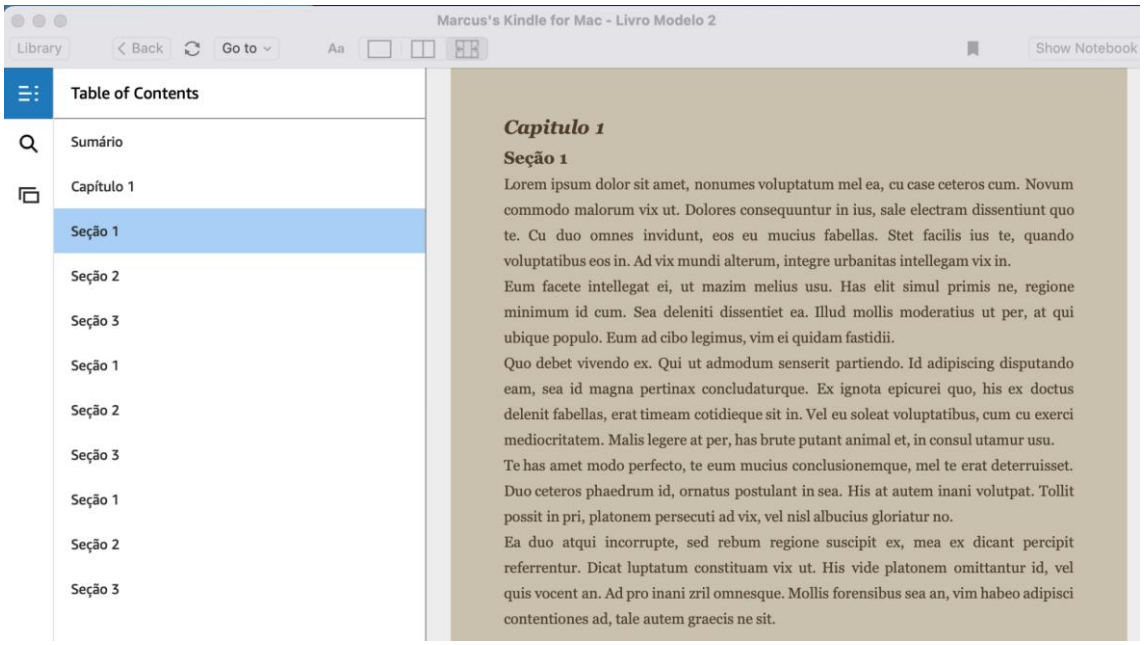
Os formatos EPUB, MOBI e AZW3 apresentaram a tabela de conteúdo corretamente no leitor do Calibre e do iBooks.



A versão convertida para o formato PDF também apresentou corretamente a tabela de conteúdo em um leitor padrão do MacOS.



Porém, ao abrir o formato MOBI no Kindle, a versão apresentada trazia erros da tabela de conteúdo, não mostrando todos os capítulos, embora trouxesse todas as seções.



Isso mostra que a conversão ocorreu da maneira correta, porém a implementação de cada leitor pode variar e apresentar o conteúdo de maneira equivocada.

5.2.7 Observação da Capa

Ao adicionar uma obra no Calibre no formato HTML compactado no formato ZIP, constatou-se que não é possível extrair a imagem da capa automaticamente do container e adicionar nas obras convertidas.

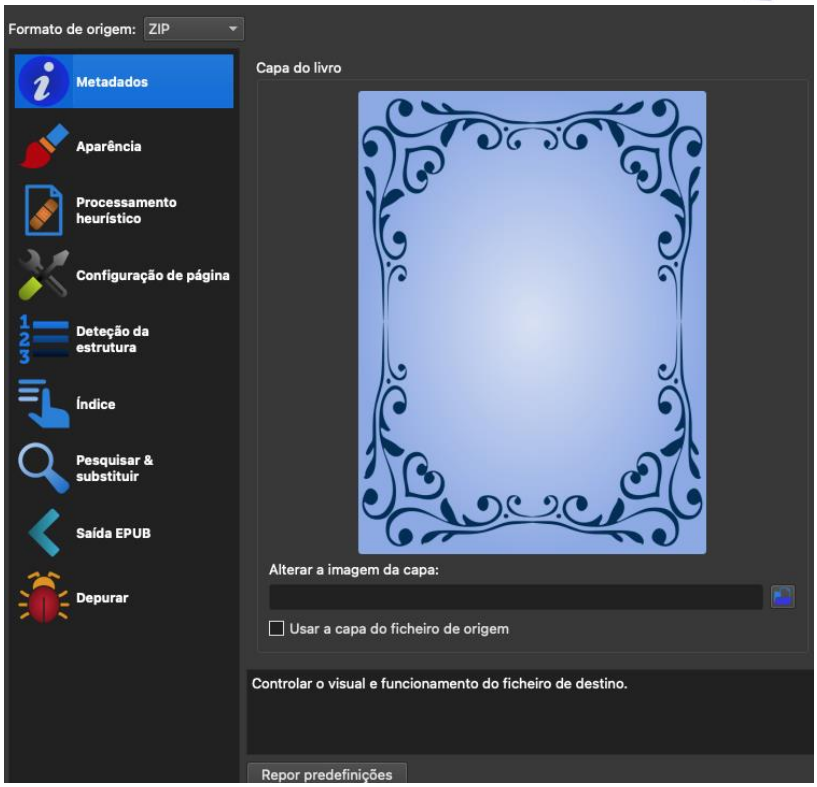
Foram realizados diversos testes adicionando a imagem da capa como um metadado tanto no index.html diretamente, quanto no arquivo OPF, seguindo todas as diretrizes descritas na normalização do arquivo OPF e incluindo os parâmetros específicos para a construção do arquivo OPF para o formato EPUB presente em https://www.w3.org/publishing/epub3/epub-packages.html#sec-cover-image.

Foram incluídos todos os metadados, implementado uma página html para apresentação da capa e adicionado ao arquivo OPF.

```
  <meta name="cover" content="cover.jpeg" />
  <x-metadata>
    <EmbeddedCover>cover.jpeg</EmbeddedCover>
  </x-metadata>
</metadata>

<manifest>
  <item id="content" href="index.html" media-type="application/xhtml+xml" />
  <item id="ncx" href="toc.ncx" media-type="application/x-dtbncx+xml" />
  <item id="global_style" href="global.css" media-type="text/css" />
  <item id="cover" properties="cover-image" href="cover.jpeg" media-type="image/jpeg" />
</manifest>

<spine toc="ncx">
  <itemref idref="content" />
</spine>

<guide>
  <reference href="index.xhtml#toc" type="toc" title="Table of Contents" />
  <reference href="index.xhtml#cap1" type="text" title="Capítulo 1" />
  <reference href="index.xhtml#cap2" type="text" title="Capítulo 2" />
  <reference href="index.xhtml#cap3" type="text" title="Capítulo 3" />
</guide>
```

Todos os testes realizados para detecção e inserção da capa do livro na obra convertida, para todos os formatos de saída, falharam e não apresentaram a capa da obra no resultado final.

Porém, o Calibre oferece uma configuração específica para alterar a capa da obra no momento da conversão. Usando a interface gráfica, a editora poderá especificar qual a capa do livro deve ser adicionada à obra resultante da conversão.



Essa mesma opção pode ser encontrada na ferramenta de conversão por linha de comando, **ebook-convert**, disponibilizada pelo Calibre, garantindo que seja possível adicionar a capa automaticamente ao efetuar a conversão por meio do sistema PNLD Digital.

Temos duas formas de executar a conversão da obra por meio da linha de comando.

1. A obra deve estar no formato compactado ZIP e a capa no formato de imagem, em arquivos separados para a execução do seguinte comando.

   $ ./ebook-convert obra.zip resultado.epub --cover cover.jpeg -d -v

2. Podemos criar um padrão para os elementos da obra e, ao receber a obra para conversão, descompactar a estrutura da obra e utilizar um comando mais completo, apontando para cada elemento da obra para realizar a conversão corretamente.

   $ ./ebook-convert index.html resultado.mobi --from-opf content.opf --cover cover.jpeg -d -v

5.2.8 Conclusão: Livro Modelo 2

Observando os resultados obtidos, verifica-se que o processo de conversão apresentou um resultado satisfatório ao manter os metadados das obras corretamente, porém observa-se a necessidade de criação de um padrão para a estrutura das obras disponibilizadas a fim de garantir maior interoperabilidade entre os diferentes formatos.

A padronização das obras deverá ser realizada tanto para nomenclatura quanto para a estrutura dos arquivos que a compõem, garantindo que, ao ser submetida ao PNLD, todas as conversões possam ser realizadas da forma correta. Assim, como pode ser visto no item 5.2.7, será importante que o sistema conheça exatamente os arquivos internos à obra para que seja possível o mapeamento correto desses arquivos no momento da conversão por meio da linha de comando.

5.3 Livro Modelo 3 - Obra com múltiplos arquivos e estilos.

Teste voltado à conversão das obras com múltiplos arquivos HTML, cada qual com diversos estilos e formatos para teste de navegação entre os conteúdos e apresentação dos estilos.

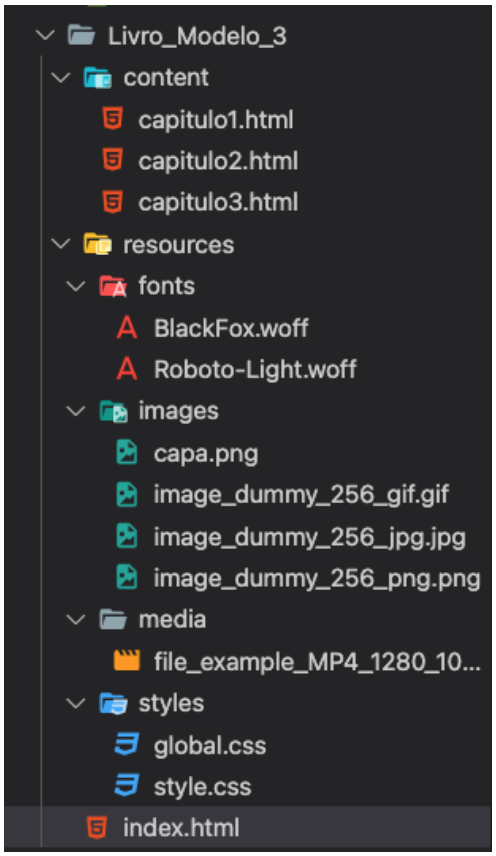
5.3.1 Preparação da obra

Para a preparação da obra "Livro Modelo 3", o projeto foi estruturado para que houvesse uma separação semântica de todo conteúdo apresentado, adicionando múltiplos arquivos HTML para separação dos capítulos e a criação de uma estrutura para todos os recursos disponíveis.

Esse modelo apresenta a uma folha de estilos diversa, composta de 2 arquivos que são adicionados ao arquivo HTML para representar as múltiplas possibilidades de diagramação de uma obra literária mais complexa.

Foram adicionados os recursos de duas fontes distintas, Roboto e BackFox, para validar a inserção de diferentes tipografias e também imagens nos formatos JPG, PNG e GIF com diferentes alinhamentos.

Por fim, a obra também contempla um recurso de vídeo para que possamos validar a renderização desse tipo de recurso nos diferentes formatos e o resultado na obra quando o leitor não suportar a renderização.



5.3.2 Objetivo

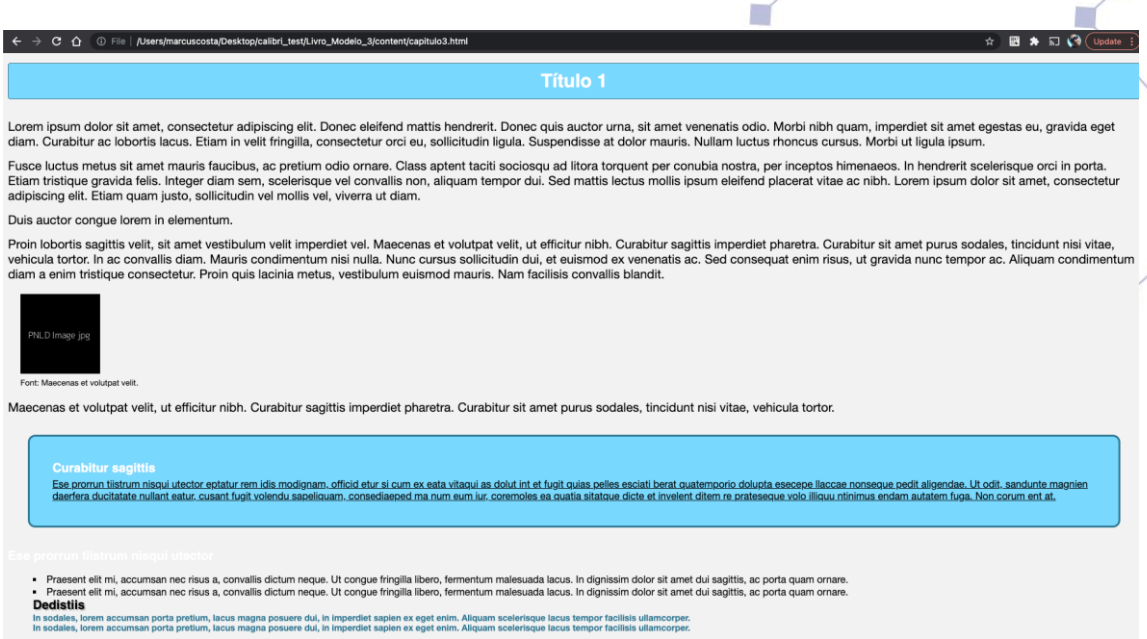
Esse cenário tem como objetivo validar a renderização de diferentes estilos aplicados ao HTML através do CSS, como bordas, fontes, cores, etc.

Como as páginas, capítulos estão dispostos em diversos arquivos em uma estrutura mais complexa de dados, deve ser verificado se o resultado final se mostra fiel ao desenvolvido em HTML e se todos os recursos disponíveis são renderizados corretamente.

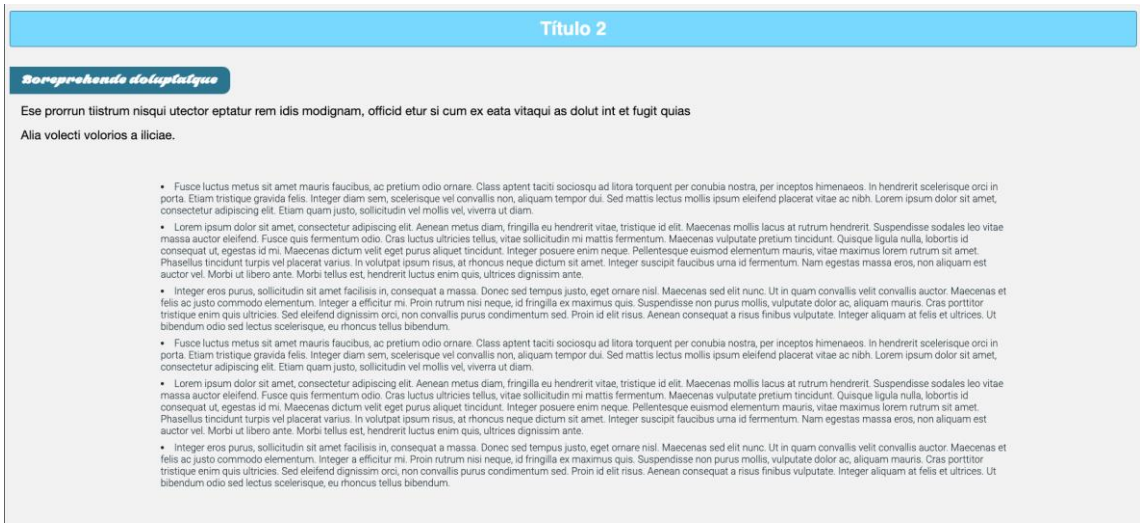
5.3.3 Apresentação da obra original - HTML

A obra original possui um sumário similar as obras anteriores, porém apontando para diferentes páginas HTML.

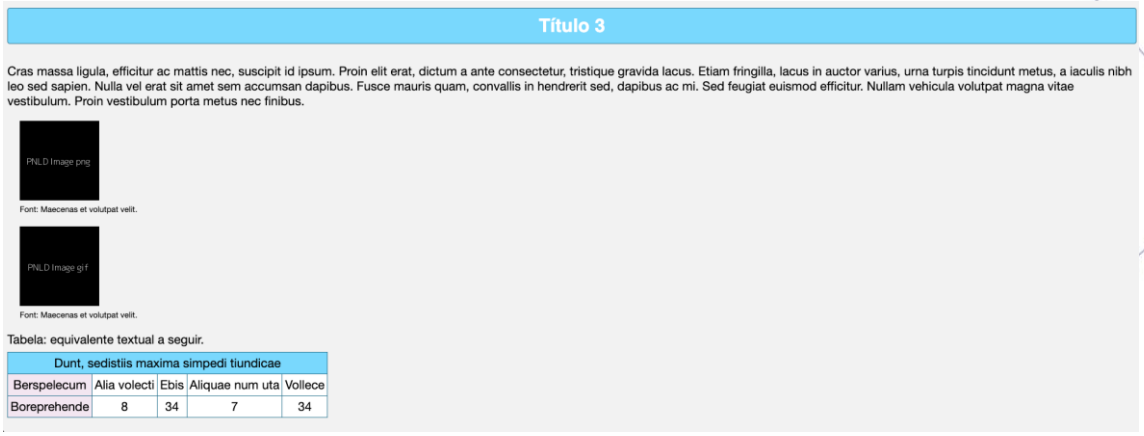
O Modelo possui alguns recursos como cores e bordas para teste de renderização.



A seção 2 de todas as páginas apresenta um *layout* com as fontes inseridas na obra, sendo o subtítulo escrito com a fonte BlackFox e o conteúdo em lista escrito com a fonte Roboto-Light.



A terceira seção das páginas apresenta as imagens no formato png e gif e uma tabela estilizada.



Título 3

Cras massa ligula, efficitur ac mattis nec, suscipit id ipsum. Proin elit erat, dictum a ante consectetur, tristique gravida lacus. Etiam fringilla, lacus in auctor varius, urna turpis tincidunt metus, a iaculis nibh leo sed sapien. Nulla vel erat sit amet sem accumsan dapibus. Fusce mauris quam, convallis in hendrerit sed, dapibus ac mi. Sed feugiat euismod efficitur. Nullam vehicula volutpat magna vitae vestibulum. Proin vestibulum porta metus nec finibus.

PNLD Image png

Font: Maecenas et volutpat velit.

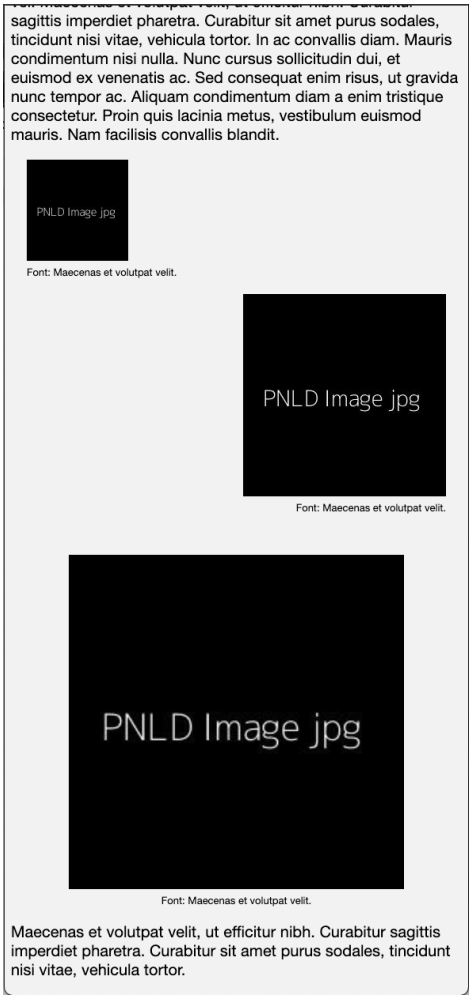
PNLD Image gif

Font: Maecenas et volutpat velit.

Tabela: equivalente textual a seguir.

Dunt, sedistiis maxima simpedi tiundicae				
Berspelecum	Alia volecti	Ebis	Aliquae num uta	Vollece
Boreprehende	8	34	7	34

Para validar o posicionamento dos elementos na diagramação da obra, foram adicionadas imagens com o posicionamento a direita, esquerda e central, sendo que a imagem apresentada no centro da página deverá ser renderizada seguindo a proporção da imagem de acordo com 80% da largura disponível.



sagittis imperdiet pharetra. Curabitur sit amet purus sodales, tincidunt nisi vitae, vehicula tortor. In ac convallis diam. Mauris condimentum nisi nulla. Nunc cursus sollicitudin dui, et euismod ex venenatis ac. Sed consequat enim risus, ut gravida nunc tempor ac. Aliquam condimentum diam a enim tristique consectetur. Proin quis lacinia metus, vestibulum euismod mauris. Nam facilisis convallis blandit.

PNLD Image jpg

Font: Maecenas et volutpat velit.

PNLD Image jpg

Font: Maecenas et volutpat velit.

PNLD Image jpg

Font: Maecenas et volutpat velit.

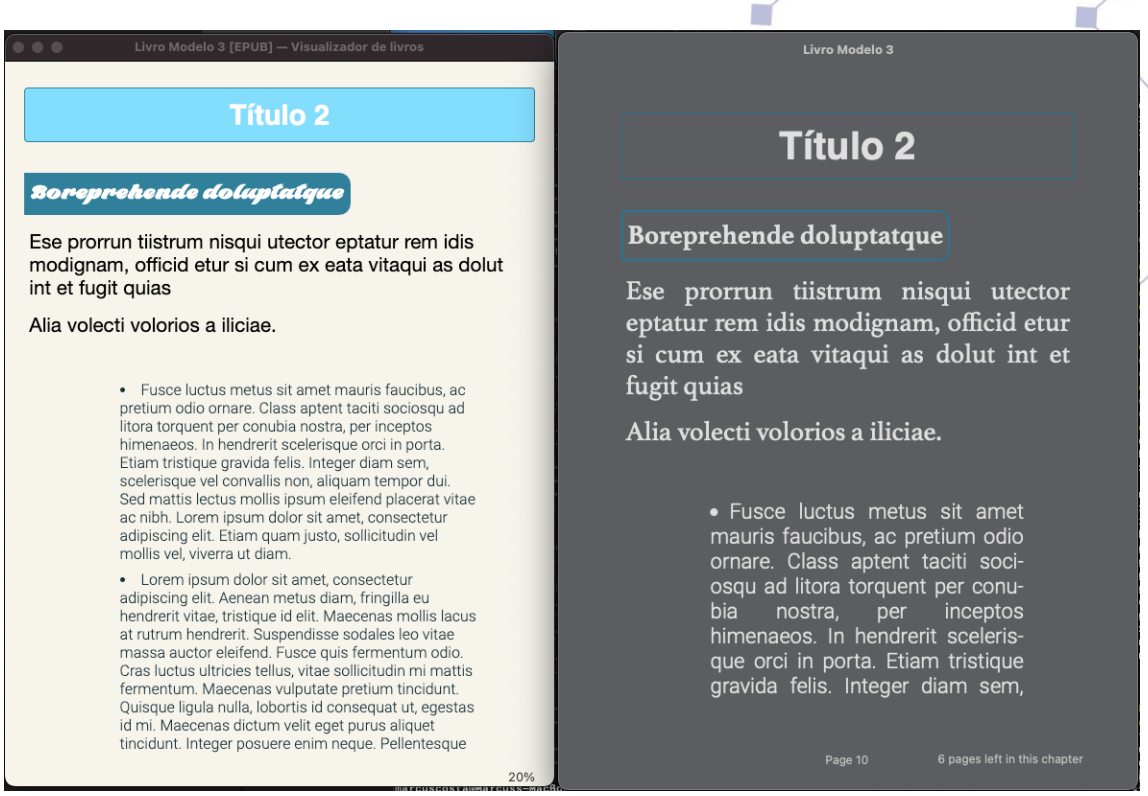
Maecenas et volutpat velit, ut efficitur nibh. Curabitur sagittis imperdiet pharetra. Curabitur sit amet purus sodales, tincidunt nisi vitae, vehicula tortor.

Finalizando a descrição da obra original no Capítulo 2, Seção 3, é preciso avaliar a renderização de um card contendo um título, descrição e um *player* de vídeo.

5.3.4 Conversão para EPUB

O formato apresentou um bom resultado de saída, tanto no Ebook-Viewer quanto no iBooks, porém é visível o impacto da implementação do leitor no resultado final obtido.

O Ebook-Viewer tem uma implementação mais completa com relação aos estilos aplicados, fazendo a renderização correta dos formatos, cores e fontes da página, porém o iBooks implementa apenas o básico na apresentação dos recursos visuais da obra, mantendo as bordas e sua espessura sem alteração das cores.



Outro ponto observado no iBooks é que houve alteração da fonte na lista, mas não para a fonte Roboto, conforme pode ser observado no Ebook-Viewer. Isso se deve ao fato de o iBooks não implementar um recurso para fontes externas, reconhecendo assim somente as fontes disponíveis no sistema.

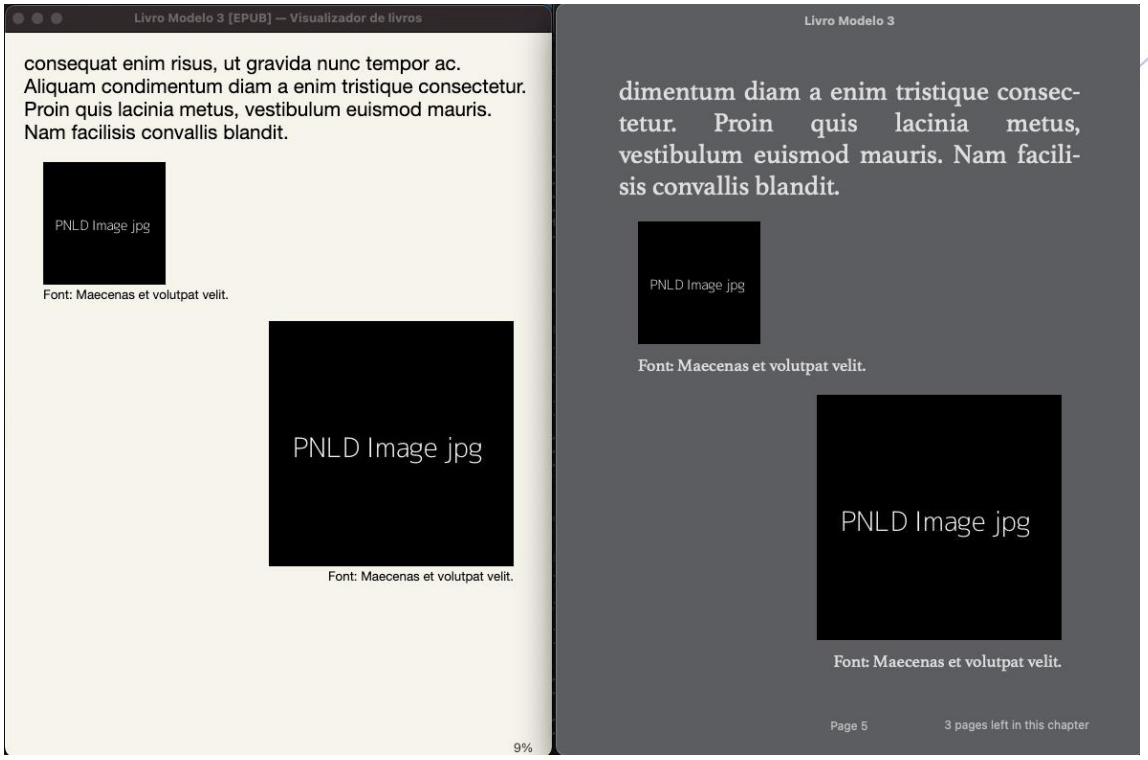
A troca da fonte aconteceu devido ao Fallback no código css na aplicação de estilo, que informa ao leitor deve usar outra fonte caso de não ser encontrada a fonte escolhida. Nesse caso, utilizando a fonte *"Helvetica Neue"* padrão para os sistemas Apple.

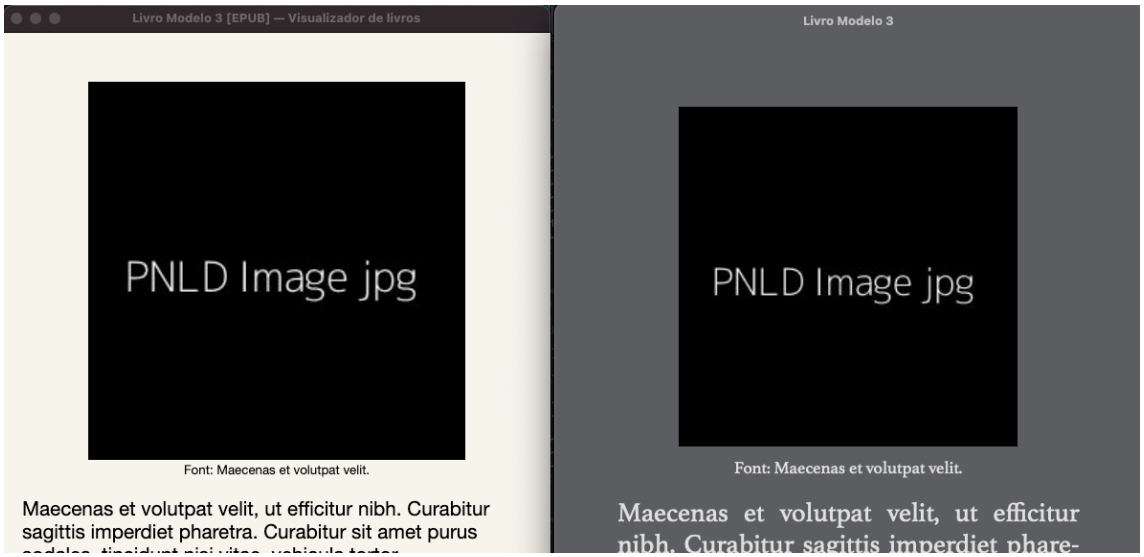
```
li.pnld_list_item {
  color: #162b33;
  font-family: "roboto", "Helvetica Neue", sans-serif;
  font-size: 16px;
```

Com base nesse evento, podemos observar a importância da aplicação de fontes como Fallback para os principais leitores do mercado, garantindo assim que a renderização da obra ocorra da melhor maneira possível.

Com relação a diagramação da obra, ambos os leitores apresentaram um resultado satisfatório, já que, de modo geral, todos os elementos e posicionamentos foram mantidos na posição

correta na página, inclusive no quesito proporcionalidade, na qual a imagem deveria manter a posição central da tela ocupando um total de 80% da largura disponível. A proporcionalidade é importante pois permite a criação de *layouts* responsivos, que se ajustam ao tamanho do equipamento do usuário, garantindo assim a melhor experiência possível com o conteúdo.

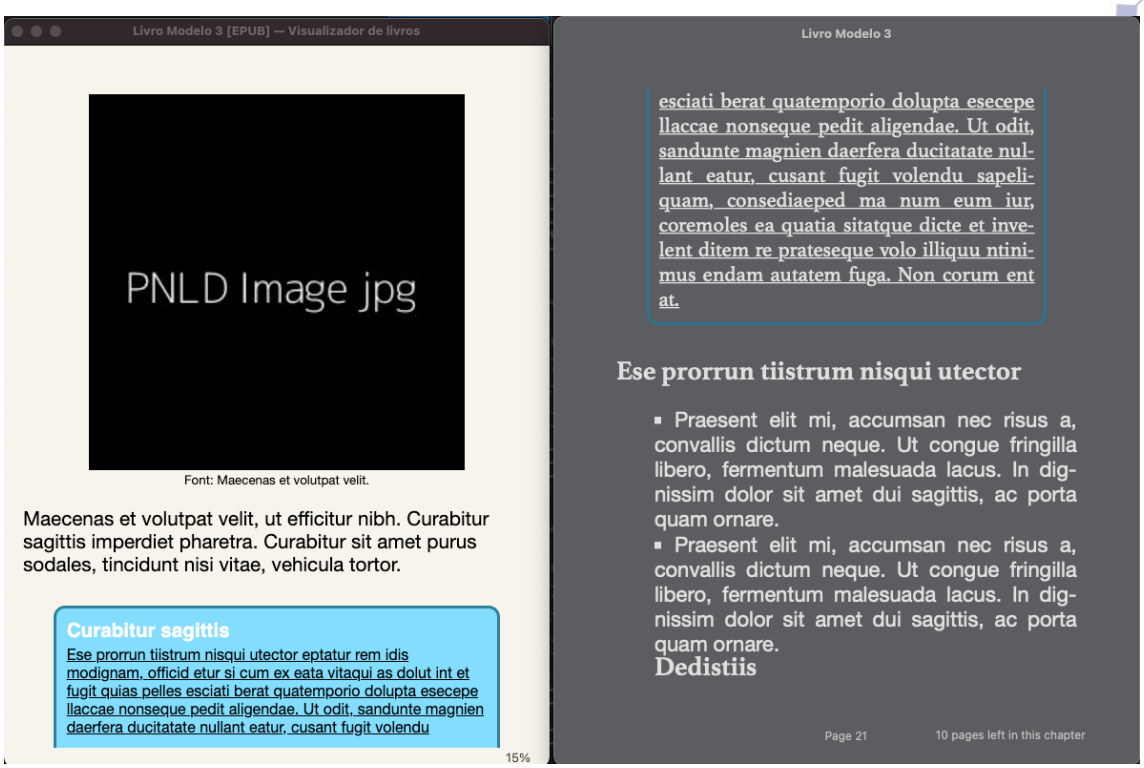




Um ponto importante a ser observado na renderização da obra está na diagramação de grandes elementos. Ao ser renderizado uma imagem o conversor automaticamente valida o tamanho da

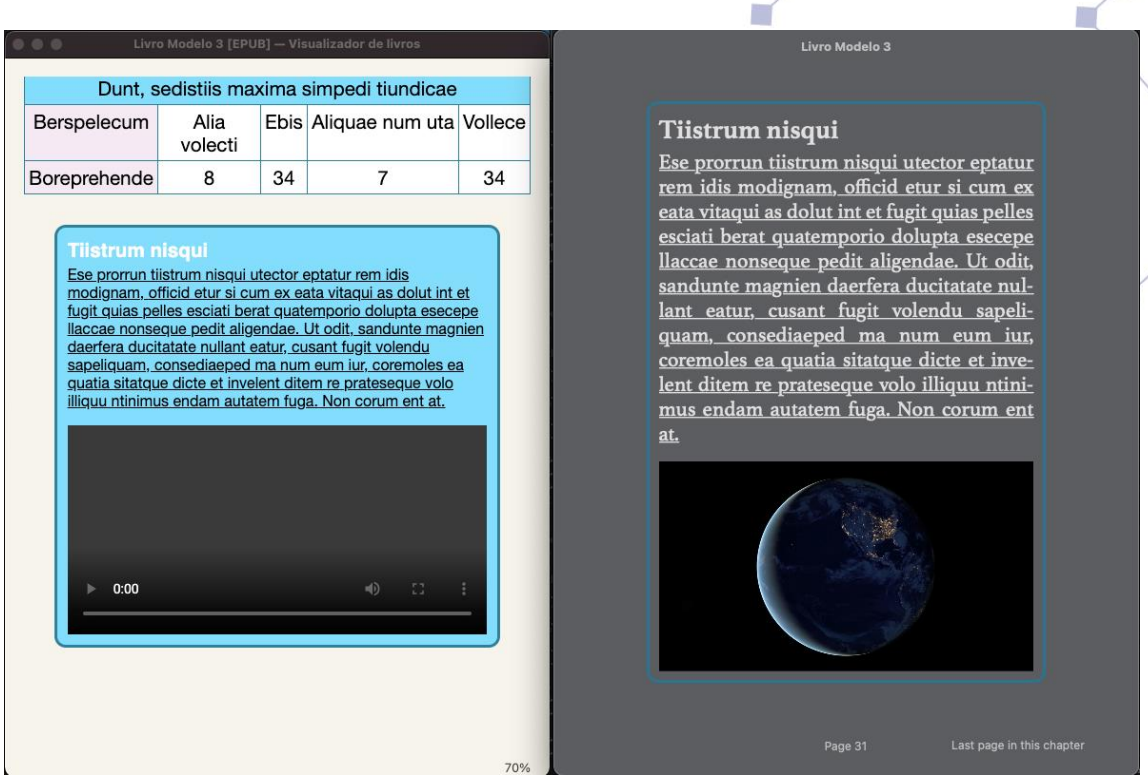
imagem para renderização e, caso o espaço disponível da tela não seja suficiente, a imagem é renderizada na página seguinte, garantindo a qualidade do conteúdo.

Porém, o mesmo não ocorre com outros elementos visuais disponíveis, sendo possível que o elemento final seja renderizado uma parte em uma tela e finalizado na outra, sempre garantindo o estilo aplicado.



Para evitar que um determinado conteúdo não seja dividido entre páginas, as editoras devem utilizar a propriedade *"page-break-before"* na folha de estilos CSS para que o leitor identifique que esse conteúdo deve ser renderizado em uma nova página ou utilizar a propriedade *"page-break-after"* para renderizar o próximo elemento em uma nova página. Essa técnica funcionou bem para a quebra de conteúdo entre páginas e visa a garantir uma melhor experiência de leitura para o usuário.

Nesse cenário, o conversor também adicionou corretamente o elemento multimídia na obra final de maneira satisfatória. Contudo, essa funcionalidade também depende do leitor para ser utilizada de forma correta. Foi observado, por exemplo, que, apesar de ambos os leitores renderizarem os elementos corretamente, tanto estilo quanto os controles de vídeo o Ebook-Viewer não foram capazes de reproduzir o elemento, enquanto o iBooks permitiu utilizar todos os recursos disponíveis para o controle da mídia.

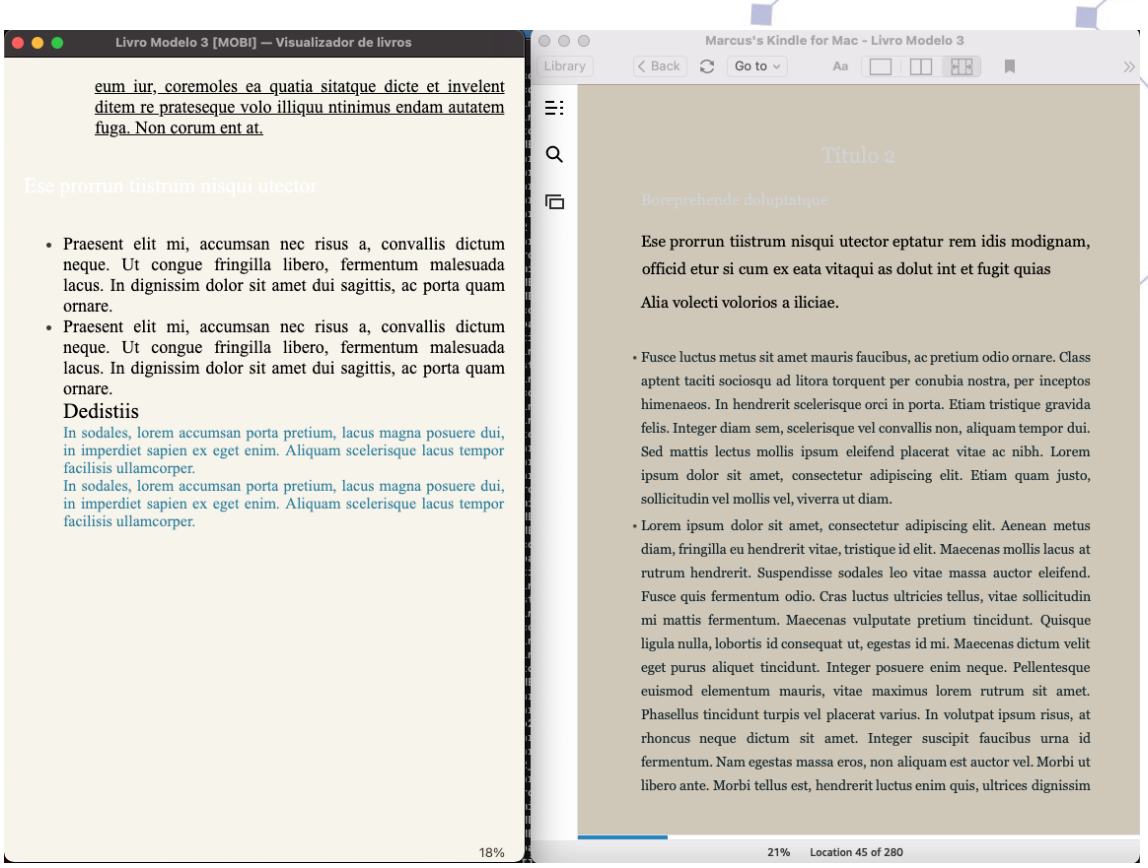


5.3.5 Conversão para MOBI e AZW3

Os formatos MOBI e AZW3 apresentam bastante diferença no teste, pois o AZW3 foi criado justamente para melhorar a aplicação de estilos que não era possível por meio do formato MOBI.

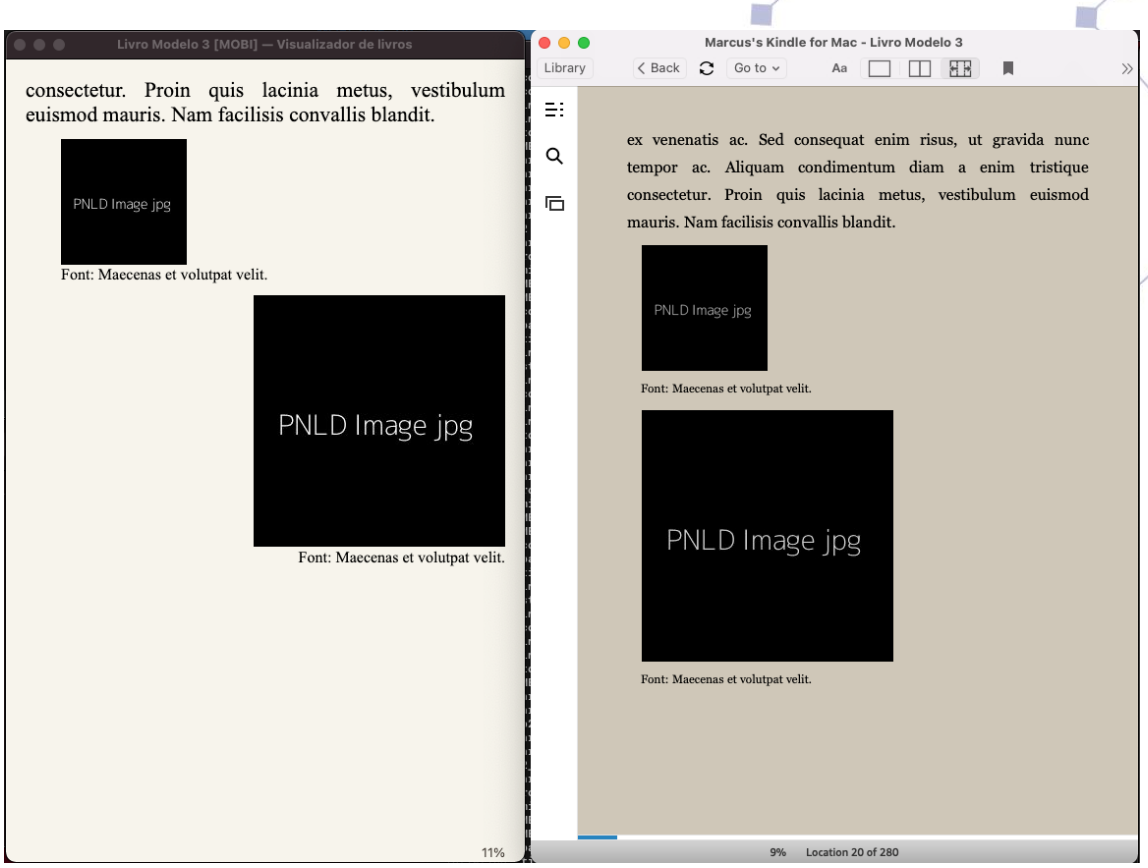
No caso específico do formato MOBI, a maioria dos estilos aplicados foram perdidos e/ou ignorados na renderização. Porém, pode ser observado que alguns pontos são referentes a conversão de formato e outros referentes ao leitor.

No caso do conteúdo, algumas informações aparentemente foram perdidas, como as cores dos títulos, que ficaram com a cor padrão em branco. Já em outros lugares, apresentou a cor azul estabelecida na folha de estilo corretamente.



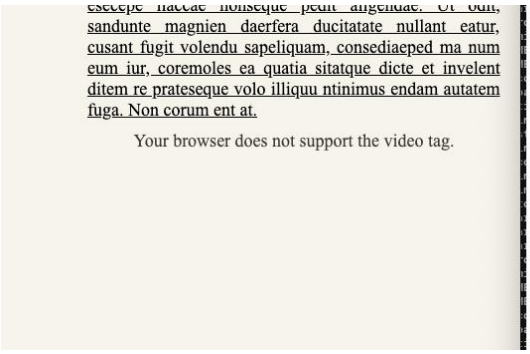
Na *guideline* disponibilizada pela Amazon em https://kindlegen.s3.amazonaws.com/AmazonKindlePublishingGuidelines.pdf, há a orientação para não ser feita a alteração da cor da fonte do corpo do texto de forma a garantir a melhor legibilidade do código. Há também a indicação casos em que o *background* tem uma cor próxima ao do texto, caso em que uma cor de *fallback* pode ser aplicada. Porém, nos testes, as cores de alguns lugares foram perdidas, sem uma lógica válida para tal e, como ambos os leitores apresentaram o mesmo problema, pode-se concluir que essa variação ocorreu diretamente na conversão da obra.

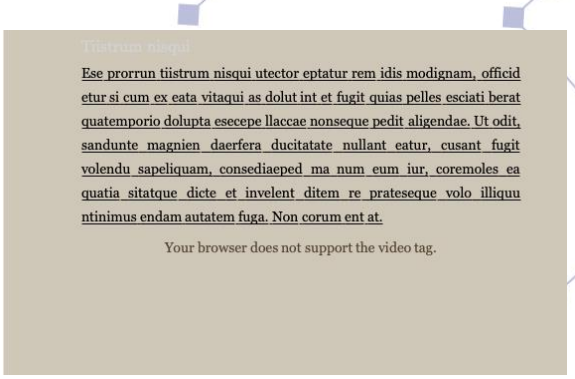
Apesar da variação na aplicação da folha de estilo, a diagramação ficou bem próximo do modelo original no quesito posicionamento dos elementos. Contudo, observou-se que esse ponto pode depender da implementação do leitor, já que no Ebook-Viewer todas as imagens foram posicionadas da forma correta, mas no Kindle isso não aconteceu.



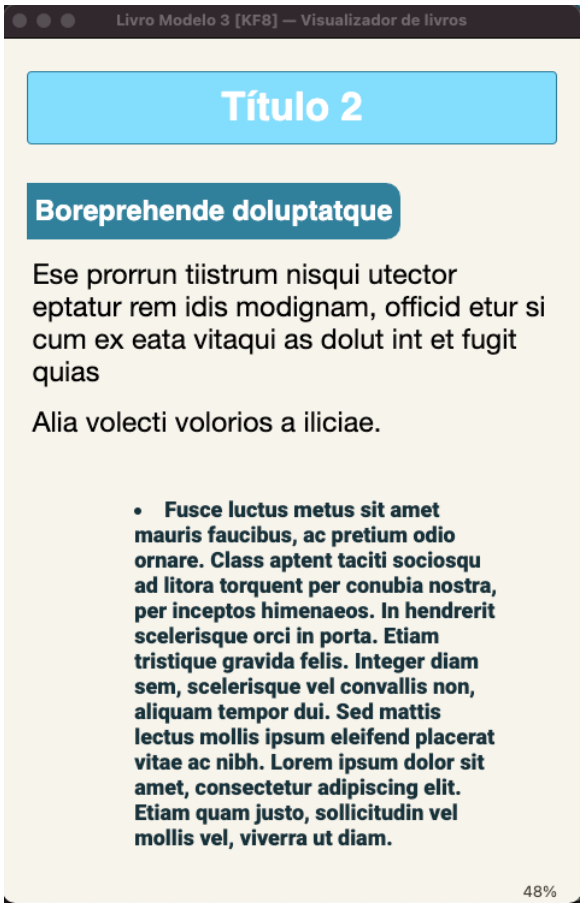
Apesar de o formato MOBI não permitir a inserção de elementos de mídia, o leitor se comportou de maneira esperada após a conversão, apresentando o texto de aviso (fallback) conforme implementado no código HTML, para leitores que não suportam a tag “video”.

```
  <div class="pnld_video_box"></div>
  <video width="100%" height="auto" controls="true">
    <source
      src="../resources/media/file_example_MP4_1280_10MG.mp4"
      type="video/mp4"
    />
    Your browser does not support the video tag.
  </video>
</div>
```

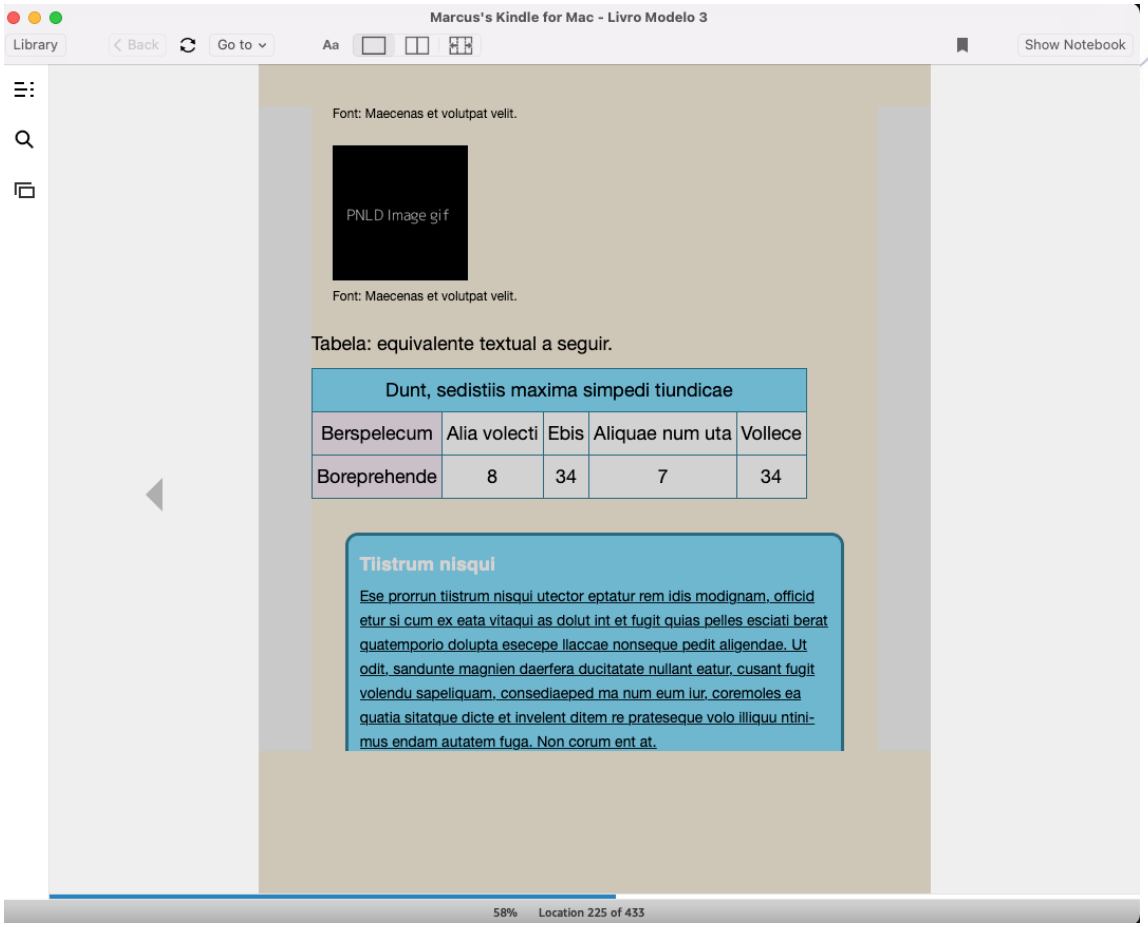




O formato AZW3, diferente do formato MOBI, permite a implementação de estilos mais livremente. Nesse foi possível obter resultados similares na renderização dos elementos comparado ao formato EPUB no Ebook-Viewer, com exceção para a alteração da fonte, que foi renderizada com uma fonte padrão. Ao avaliar a documentação da *guideline* de publicação do Kindle, observou-se que os formatos disponibilizados pela Amazon, AZW3 e MOBI só permitem a utilização de fontes customizadas nos formatos Open Type (OTF) ou True Type (TTF).



É importante observar que, apesar do formato AZW3 ter apresentado um resultado satisfatório no leitor Ebook-Viewer, ele não foi renderizado de maneira satisfatória no Kindle, apresentando diversas falhas na renderização dos elementos desde o *background* a alguns cortes secos em algum ponto do *layout*.



5.3.6 Conversão para PDF

O formato apresentou um bom resultado de saída mantendo todos os estilos estabelecidos no *layout,* como cores, fontes, bordas, etc., e também a diagramação correta referentes aos posicionamentos dos elementos e proporcionalidade.

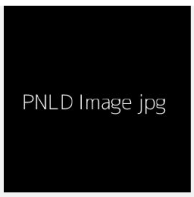
Título 2

Boreprehende doluptatque

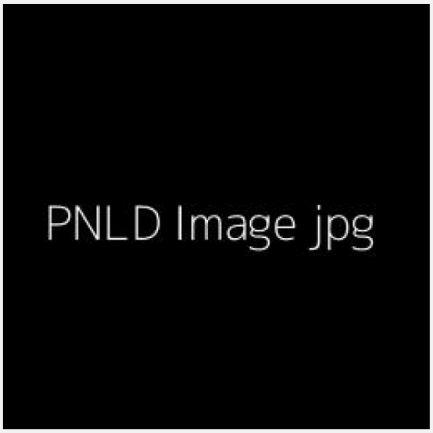
Ese prorrun tiistrum nisqui utector eptatur rem idis modignam, officid etur si cum ex eata vitaqui as dolut int et fugit quias

Alia volecti volorios a iliciae.

- Fusce luctus metus sit amet mauris faucibus, ac pretium odio ornare. Class aptent taciti sociosqu ad litora torquent per conubia nostra, per inceptos himenaeos. In hendrerit scelerisque orci in porta. Etiam tristique gravida felis. Integer diam sem, scelerisque vel convallis non, aliquam tempor dui. Sed mattis lectus mollis ipsum eleifend placerat vitae ac nibh. Lorem ipsum dolor sit amet, consectetur adipiscing elit. Etiam quam justo, sollicitudin vel mollis vel, viverra ut diam.
- Lorem ipsum dolor sit amet, consectetur adipiscing elit. Aenean metus diam, fringilla eu hendrerit vitae, tristique id elit.

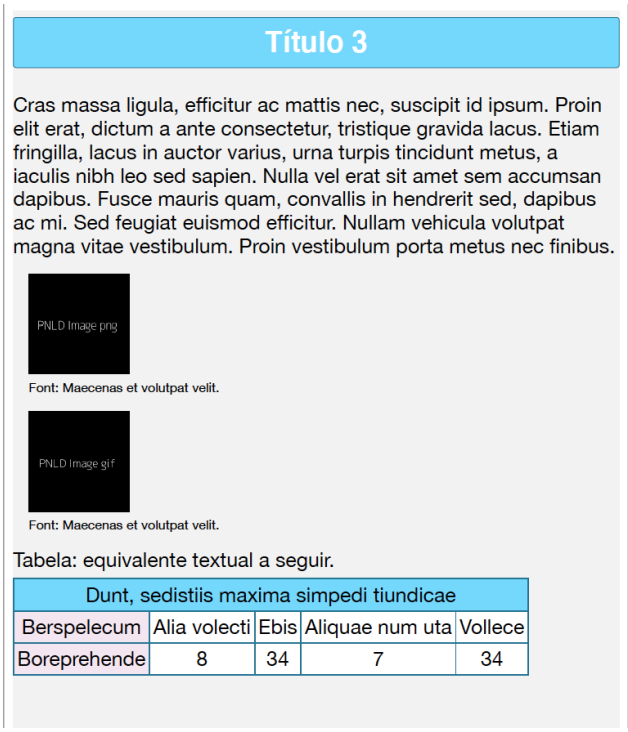


Font: Maecenas et volutpat velit.



Font: Maecenas et volutpat velit.

Maecenas et volutpat velit, ut efficitur nibh. Curabitur sagittis imperdiet pharetra. Curabitur sit amet purus sodales, tincidunt nisi



Título 3

Cras massa ligula, efficitur ac mattis nec, suscipit id ipsum. Proin elit erat, dictum a ante consectetur, tristique gravida lacus. Etiam fringilla, lacus in auctor varius, urna turpis tincidunt metus, a iaculis nibh leo sed sapien. Nulla vel erat sit amet sem accumsan dapibus. Fusce mauris quam, convallis in hendrerit sed, dapibus ac mi. Sed feugiat euismod efficitur. Nullam vehicula volutpat magna vitae vestibulum. Proin vestibulum porta metus nec finibus.

PNLD Image.png

Font: Maecenas et volutpat velit.

PNLD Image.gif

Font: Maecenas et volutpat velit.

Tabela: equivalente textual a seguir.

Dunt, sedistiis maxima simpedi tiundicae				
Berspelecum	Alia volecti	Ebis	Aliquae num uta	Vollece
Boreprehende	8	34	7	34

Ao se observar os resultados obtidos, dois pontos chamam a atenção.

1. A margem da página fica visível na conversão da obra e esse pode não ser um efeito interessante. Esse é um ponto configurável no Calibre no momento de converter o formato. Para isso, pode-se definir um tamanho de página e das margens da forma livre para se obter o melhor resultado para a obra. Outra opção que pode ser utilizada é trabalhar com a cor de fundo da obra e controlar todas as margens diretamente por meio da folha de estilos CSS.

2. Corte do conteúdo entre as páginas.

O conversor não valida se o bloco a ser renderizado pode ser comportado no espaço restante da folha, o que ocasionalmente pode apresentar uma falha na renderização do layout. O mesmo efeito acontece com as outras versões exportadas.

Para evitar que um determinado conteúdo não seja dividido entre páginas, as editoras devem utilizar a propriedade *"page-break-before"* na folha de estilos CSS para que o leitor identifique que esse conteúdo deve ser renderizado em uma nova página ou utilizar a propriedade *"page-break-after"* para renderizar o próximo elemento em uma nova página. Essa técnica funcionou bem para a quebra de conteúdo entre páginas e visa a garantir uma melhor experiência de leitura para o usuário.

5.3.7 Conclusão: Livro Modelo 3

Observando os resultados obtidos em cada um dos leitores, exceto no formato MOBI, o processo de conversão apresenta um resultado satisfatório para a implementação das obras.

O formato MOBI, da Amazon, não está apto para aplicação de estilos complexos, devendo as editoras tomarem cuidado na formatação da obra, caso esse seja um formato de saída válido. Na implementação do formato, a editora deverá fazer a diagramação do conteúdo normalmente, sendo aconselhável que se mantenha as cores padrão do leitor para garantir uma boa experiência.

Outro ponto de atenção é com relação ao corte do conteúdo entre páginas. Nos formatos digitais esse ponto pode ser bem aceito pelos usuários devido a dinâmica de troca de tamanho de fonte para leitura, o que pode fazer com que o conteúdo mude de posição e não apresenta necessariamente uma falha. Porém, para as obras convertidas para o formato PDF, pode ser necessário uma atenção especial no momento da diagramação de forma a garantir que o livro seja apresentado da melhor maneira possível.

5.4 Livro Modelo 4 - Obra com tags de acessibilidade

Teste voltado para a conversão das obras com tags de acessibilidade, definidas pela *Guideline* de acessibilidade para conteúdo Web (WCAG).

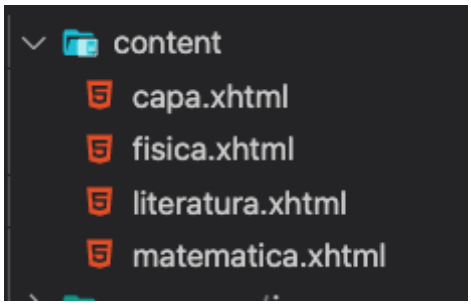
Diferentemente dos modelos apresentados anteriormente, que visavam a uma maior interoperabilidade entre os formatos, o Livro Modelo 4 busca a melhor relação de acessibilidade entre as obras. Por esse motivo, os testes não foram realizados na transformação para o EPUB na versão 2.

O EPUB versão 2 não comporta diversas tags do HTML5 que tem por objetivo melhorar a semântica do código da obra, dando a relevância correta ao conteúdo apresentado, que são utilizados pelos leitores de telas mais modernas.

5.4.1 Preparação da obra

Para preparação da obra "Livro Modelo 4" foram criados conteúdos fictícios para representar livros didáticos reais, em texto legível para validar a leitura do conteúdo por meio de sistemas especializados, como o VoiceOver.

Para teste de navegação, foram criados conteúdos, em arquivos separados, e ordenados por meio de um sumário na página principal da obra (*index.xhtml*) com diferentes níveis.




```
<nav role="doc-toc" id="toc">
  <h1>Sumário</h1>
  <ol>
    <li><a href="content/capa.xhtml">Capa</a></li>
    <li>
      <a href="content/literatura.xhtml">Literatura</a>
      <ol>
        <li>
          <a href="content/capitulo1.html#presentation">Apresentação</a>
        </li>
        <li><a href="content/capitulo1.html#note">Nota Informativa</a></li>
        <li><a href="content/capitulo1.html#biography">Biografia</a></li>
        <li><a href="content/capitulo1.html#capitulo1">Capitulo1</a></li>
      </ol>
    </li>
    <li>
      <a href="content/matematica.xhtml">Matemática</a>
    </li>
    <li>
      <a href="content/fisica.xhtml">Física</a>
    </li>
  </ol>
</nav>
```

Como observado nos testes anteriores, é importante a criação da tabela de conteúdo - TOC também através do arquivo ncx. De um modo geral, a maioria dos leitores criam uma tabela de conteúdo automaticamente por meio da leitura desse arquivo de maneira a garantir que todos os *links* de navegação serão criados de forma correta e que haverá a melhor experiência de navegação usando os sistemas especializados em acessibilidade.

No arquivo que descreve todo o conteúdo disponível na obra (content.opf), é importante descrever os recursos de acessibilidade utilizados na obra que serão utilizados pelos leitores de conteúdo para garantir uma melhor experiência do usuário.

```
<meta property="schema:accessibilityFeature">structuralNavigation</meta>
<meta property="schema:accessibilityFeature">tableOfContents</meta>
<meta property="schema:accessibilityFeature">MathML</meta>
<meta property="schema:accessibilityFeature">describedMath</meta>
<meta property="schema:accessibilityAPI">ARIA</meta>
</metadata>
```

É importante adicionar as mesmas propriedades de acessibilidade no conteúdo gerado em HTML para que, na utilização da obra por meio do browser, diretamente em HTML5, os recursos sejam identificados corretamente.

```
<body lang="pt-BR" xml:lang="pt-BR">
  <div itemscope="" itemtype="https://schema.org/Book">
    <meta itemprop="accessibilityFeature" content="largePrint/CSSEnabled" />
    <meta itemprop="accessibilityFeature" content="highContrast/CSSEnabled" />
    <meta itemprop="accessibilityFeature" content="resizeText/CSSEnabled" />
    <meta itemprop="accessibilityFeature" content="displayTransformability" />
    <meta itemprop="accessibilityFeature" content="longDescription" />
    <meta itemprop="accessibilityFeature" content="alternativeText" />
    <meta itemprop="accessibilityControl" content="fullKeyboardControl" />
  </div>
```

E adicionar metas específicas de cada conteúdo diretamente no elemento ao qual ele pertence, garantindo que todos os leitores possam ler essas informações, seja por meio do arquivo de conteúdo ou diretamente no HTML..

```
<div
  role="text|math"
  aria-label="Abrir parêntese 50 mais 100, fechar parêntese. Menos, abrir parêntese 50 mais 50 mais 100, fechar parênteses."
  aria-describedby="forces02"
>
  <meta property="accessibilityFeature" content="MathML" />
  <meta property="accessibilityFeature" content="describedMath" />
  <math xmlns="http://www.w3.org/1998/Math/MathML">
    <mo>(</mo>
    <mn>50</mn>
    <mo>+</mo>
    <mn>100</mn>
    <mo>)</mo>
```

Outro recurso importante para uma experiência acessível, é a implementação da tag <guide>, anteriormente conhecida como <tour>, que pode possibilitar uma melhor experiência na navegação do conteúdo da obra, conforme a documentação para o desenvolvimento do arquivo OPF disposta em: http://idpf.org/epub/20/spec/OPF_2.0.1_draft.htm#TOC2.6.

```
<guide>
  <reference href="index.xhtml#toc" type="toc" title="Table of Contents" />
  <reference href="content/literatura.xhtml" type="text" title="Literatura" />
  <reference href="content/matematica.xhtml" type="text" title="Matemática" />
  <reference href="content/fisica.xhtml" type="text" title="Física" />
</guide>
```

O conteúdo da obra foi separado em três arquivos para testes diferentes entre eles, sendo literatura, matemática e física.

No arquivo de literatura, foi adicionado o início da obra Memórias Póstumas de Brás Cubas, separando o conteúdo em diversas seções, com conteúdo em português e uma nota em inglês, tags semânticas do HTML5 como *section* e *article* e imagens com o intuito de testar a navegação da obra e a leitura por meio de ferramentas especializadas.

Memórias Póstumas de Brás Cubas (1881)

Machado de Assis, Joaquim Maria (1839-1908)

Fonte digital:

Ministério da Cultura

Fundação BIBLIOTECA NACIONAL

Departamento Nacional do Livro

Biblioteca Nacional

Copyright

©2000,2006 Joaquim Maria Machado de Assis

Memórias Póstumas de Brás Cubas

Machado de Assis

Nota Informativa

Utilizaram-se, para a revisão e fixação do texto de Memórias Póstumas de Brás Cubas, as seguintes publicações do romance:

1. Machado de Assis, Obra Completa, Volume I — Romances — Rio de Janeiro, Editora José Aguilar Ltda., 1962 — Organização de Afrânio Coutinho.
2. Machado de Assis, Memórias Póstumas de Brás Cubas, Biblioteca Folha — Ediouro Publicações S.A. Introdução e notas de Ivan Cavalcanti Proença — 1995.

Nos casos de dúvidas, foram consultadas a edição princeps e a quarta edição do romance, ambos existentes na Fundação Biblioteca Nacional.

Seguindo a tradição das edições anteriores, foram modernizadas a ortografia, os nomes próprios e os ditongos em ou, transformados em oi, conforme a lição atual.

Memórias póstumas de Brás Cubas é um romance fundamental na obra de Machado de Assis, não apenas porque inaugura o estilo psicológico no romance brasileiro como também implica a mudança qualitativa na escrita do autor.

Informative Note

For the revision and fixation of the text of Posthumous Memoirs of Brás Cubas, the following publications of the novel were used:

1. Machado de Assis, Obra Completa, Volume I — Romances — Rio de Janeiro, Editora José Aguilar Ltda., 1962 — Organização de Afrânio Coutinho.
2. Machado de Assis, Memórias Póstumas de Brás Cubas, Biblioteca Folha — Ediouro Publicações S.A. Introdução e notas de Ivan Cavalcanti Proença — 1995.

In cases of doubt, the princeps edition and the fourth edition of the novel were consulted, both existing in the National Library Foundation.

Following the tradition of previous editions, the spelling, proper names and diphthongs were modernized in or, transformed into oi, according to the current lesson.

Posthumous Memoirs of Brás Cubas is a fundamental novel in the work of Machado de Assis, not only because it inaugurates the psychological style in the Brazilian novel, but also implies a qualitative change in the author's writing.

Marcia Cavalcanti

BIOGRAFIA

JOAQUIM MARIA MACHADO DE ASSIS nasceu no Rio de Janeiro, em 21 de junho de 1839 e faleceu, nesta mesma cidade, em 29 de agosto de 1908.

Mestiço, de origem humilde — filho de um pintor de paredes e de uma imigrante açoriana — autodidata, sofrendo a inevitáveis restrições de sua descendência social, Machado de Assis, tendo perdido cedo sua genitora, foi criado por uma madrasta, que o cuidou

Sua produção abrange a poesia, o conto, o ensaio, o teatro, o publicismo, a crônica, mas foi no romance, sobretudo os cinco textos após 1881, que ele atingiu os maiores níveis de criatividade e elaboração formal.

É o autor brasileiro de mais vasta fortuna crítica, não apenas entre os críticos nacionais mas também entre os estrangeiros que o consideram um caso raro de artista e intelectual.

No arquivo de matemática, foram adicionados tags aria para leitura de fórmulas matemáticas com o intuito de testar a apresentação das fórmulas e sua leitura, assim como validar as tags aria após a conversão da obra pelo Calibre.

Livro modelo para calculo

Adicionando forcas

As forças podem ser adicionadas e subtraídas para dar uma força resultante. Para entender os fundamentos do cálculo da força, consideremos um jogo de cabo de guerra.

Os jogadores de duas equipes formam grupos em extremidades opostas de uma corda e puxam-se um contra o outro até que uma equipe vença puxando a outra equipe sobre a linha central. Quando a partida começa, cada equipe exerce força em direções opostas, e a equipe com a maior força combinada vencerá porque a força resultante está indo em sua direção.

Exemplo:

A equipe esquerda escolhe dois membros capazes de puxar 50 N e 100 N para a esquerda. A equipe certa seleciona três membros que podem puxar 50 N, 50 N, e 100 N para a direita.

Para descobrir qual equipe vencerá, calcule a força resultante somando as forças de cada equipe.

Para a equipe esquerda, a força total é de 50 N + 100 N = 150 N.

Para a equipe direita, a força total é de 50 N + 50 N + 100 N = 200 N.

(50 + 100) - (50 + 50 + 100)
= (150) - (200)
= -50 N

A força é 150 N menos 200 N (nós substraímos 200 N porque o time direito esta puxando na direção *oposta* a direção do time esquerdo). Portanto, a força total é -50 N, o que significa que a equipe direita ganha.

```
<div
  role="text|math"
  aria-label="É igual a, abrir parênteses 150, fechar parênteses. Menos, abrir parênteses 200, fechar parênteses."
  aria-describedby="forces02"
>
  <meta property="accessibilityFeature" content="MathML" />
  <meta property="accessibilityFeature" content="describedMath" />
  <math xmlns="http://www.w3.org/1998/Math/MathML">
    <mo>=</mo>
    <mo>(</mo>
    <mn>150</mn>
    <mo>)</mo>
    <mo>−</mo>
    <mo>(</mo>
    <mn>200</mn>
    <mo>)</mo>
  </math>
</div>
```

No arquivo da obra de física, foi utilizada uma tabela com o intuito de validar a leitura e navegação nesse modelo de conteúdo.

Livro modelo para fisica

Sistema Internacional de Unidades

O Sistema Internacional de Unidades, abreviado pela sigla SI, é um conjunto de unidades de medidas correspondentes às grandezas físicas fundamentais e suas derivações. O SI representou uma evolução do sistema métrico quando estabelecido em 1960, durante a Conferência Geral de Pesos e Medidas (CGPM), na França.

O Sistema Internacional de Unidades é completamente escrito sobre sete unidades de medida básicas, baseadas nas grandezas físicas fundamentais: comprimento, tempo, massa, corrente elétrica, temperatura termodinâmica, quantidade de matéria, e intensidade luminosa.

Definição do Sistema Internacional de Unidades
Tabela que mostra como foi realizada a definição de cada tipo de unidade do sistema métrico internacional

Grandeza	Unidade	Símbolo	Definição moderna
Comprimento	metro	m	O metro é definido como o espaço percorrido pela luz (no vácuo) em uma fração de 1/299.792.458 s.
Tempo	segundo	s	O segundo equivale a 9.192.631.770 transições hiperfinas de energia de um átomo de Césio.
Massa	quilograma	kg	Atualmente o quilograma passou a ser baseado na constante de Planck, igual a 6,62607015.10-34 J.s.
Corrente elétrica	ampére	A	O ampére é igual à passagem de 1,602176634.1019 cargas elementares por segundo, corresponde à corrente que produz uma força de 2.10-7 N entre dois fios condutores paralelos, espaçados em 1 m.
Temperatura termodinâmica	kelvin	K	Recentemente, a temperatura termodinâmica passou a ser medida em termos da constante de Boltzmann, de módulo igual a 1,380649.1023 J.s. Antigamente, era relacionada com o ponto triplo da água.
Quantidade de matéria	mol	mol	O mol é definido em termos do número de Avogadro, que define como 6,02214076.1023 o número de partículas contidas em um mol.
Intensidade luminosa	candela	cd	A intensidade luminosa é baseada em uma frequência monocromática de luz igual a 540.1012 Hz.

Além dessas unidades básicas, existem outras 22 unidades derivadas, como o newton (N = kg.m/s²), o joule (kg.m²/s2) e o coulomb (A.s). Para cada unidade do SI, básica ou derivada, pode-se aplicar prefixos de unidade. Ao todo, existem 20 prefixos de unidade, mostrados nesta tabela

Para a criação do teste de leitura e navegação da tabela, seguimos, inicialmente, o material disponibilizado no portal do Governo Digital, que define as boas práticas na construção de tabelas acessíveis, disponível em: https://www.gov.br/governodigital/pt-br/acessibilidade-digital/tabelas-acessiveis.pdf

É aconselhável a utilização da propriedade “summary” para a adição de informações mais ricas sobre o conteúdo da tabela a serem lidas pelos sistemas especializados. Nesse caso, ao consultar os tutoriais de acessibilidade do W3 disponível em https://www.w3.org/WAI/tutorials/tables/caption-summary/, foi observado que essa abordagem foi depreciada no HTML5.

Approach 4: Using the `summary` attribute SHARE

Note: The `summary` attribute is deprecated in HTML5.

In this approach, the summary text is in the `summary` attribute of the table. Such summaries are not displayed visually. They are commonly only available to screen reader users.

Nesse caso, optou-se por utilizar nos testes a abordagem 3, disponível no mesmo *link*, com a utilização da tag *<figure>* para agrupamento das informações em conjunto com as propriedades aria, garantindo um teste mais completo relacionado não somente a tabela, mas à tag *<figure>,* muito utilizada também na descrição de imagens nas obras.

Approach 3: Using the `<figure>` element to mark up a table summary SHARE

In this approach, the table is wrapped in a `<figure>` element. The `<figcaption>` element contains the caption and summary text.

Screen reader users navigating in "Tables Mode" are usually unable to identify a table by a caption applied like this. The caption part of the `<figcaption>` element can be explicitly associated to the table by using the `aria-labelledby` attribute and the summary part by using the `aria-describedby` attributes. Note that this could lead to the caption and summary being read out multiple times.

5.4.2 Objetivo

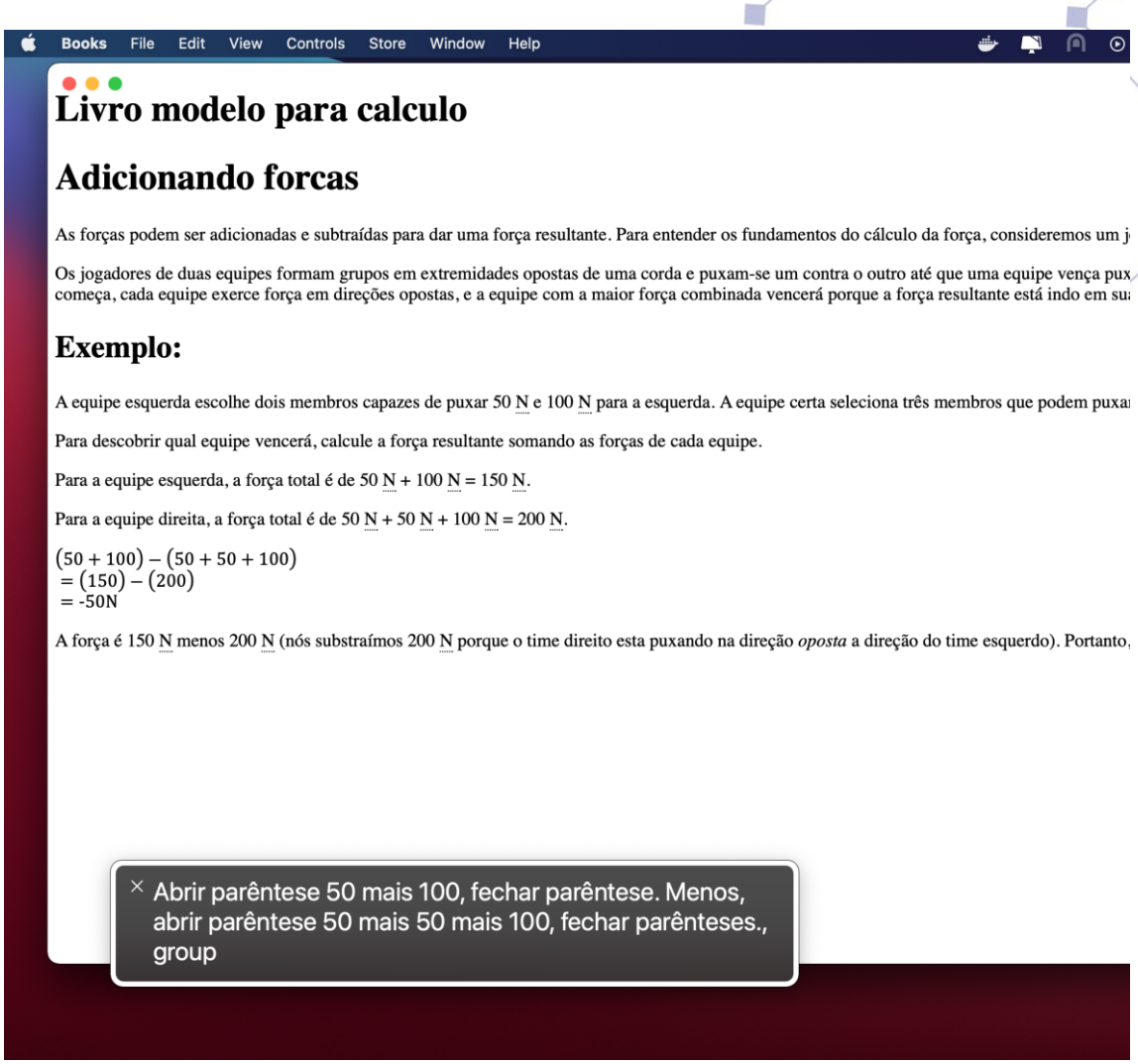
Esse teste tem como objetivo validar o suporte das tags de acessibilidade pelo conversor do Calibre, validando se o conteúdo gerado mantém a configuração e verificando qual o comportamento dos leitores e demais ferramentas de acessibilidade com relação a obra convertida.

5.4.3 Teste de leitura e navegação

Após a conversão da obra utilizando o Calibre, observamos que o código gerado, no formato EPUB3, estava de acordo com o código da obra original. Nesse caso, todas as nuances de acessibilidade adicionadas no HTML foram mantidas e são passíveis de interpretação pelos sistemas especializados.

Nos testes com o VoiceOver (MacOS) e NVDA (Windows), a navegação na obra através dos *links* utilizando o teclado e a leitura do conteúdo ocorreu de forma satisfatória, sendo preservada a ordem de leitura de acordo com a ordem dos componentes no DOM.

As variantes que se utilizam das propriedades Aria, também foram lidas de forma correta, variando apenas de acordo com o leitor e suas implementações, já que, dependendo da implementação e/ configuração do leitor os dados, são lidos de forma diferentes. Por exemplo, utilizando o sistema operacional MacOS no idioma inglês e o VoiceOver, as tags de título de *heading (h1, h2, etc.)* são faladas em inglês, porém o conteúdo é falado em português de acordo com a marcação do HTML, fazendo com que o VoiceOver se perca às vezes na leitura do idioma correto ou lendo o conteúdo nos dois idiomas.



Foi encontrada uma discrepância entre o arquivo OPF, que descreve o conteúdo original e o arquivo OPF, gerado após a conversão. O Calibre, ao realizar a conversão do formato, faz um merge entre os dados originais e os dados gerados pelo conversor, mas foi observado que o conversor descartou os metadados que foram adicionados para descrever os recursos de acessibilidade utilizados, fazendo com que no arquivo OPF final disponibilizado no formato EPUB3 não tenha esse descritivo.

Metadados originais:

```
<package version="2.0" xmlns="http://www.idpf.org/2007/opf" unique-identifier="BookId">
  <metadata xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:opf="http://www.idpf.org/2007/opf">
    <dc:title>Livro Modelo 4</dc:title>
    <dc:language>pt-br</dc:language>

    <dc:creator opf:role="aut" opf:file-as="PNLD, Autor">Autor PNLD</dc:creator>
    <dc:publisher>Editora PNLD</dc:publisher>
    <dc:date>2021-08-03</dc:date>

    <dc:description>Esse é um livro exemplo para teste do Calibre</dc:description>

    <dc:subject>Exemplo</dc:subject>
    <dc:subject>Calibre</dc:subject>
    <dc:subject>Conversão</dc:subject>

    <dc:rights>Todos os direitos reservados</dc:rights>

    <meta name="cover" content="cover.jpeg" />
    <x-metadata>
      <EmbeddedCover>cover.jpeg</EmbeddedCover>
    </x-metadata>

    <meta property="schema:accessibilityFeature">structuralNavigation</meta>
    <meta property="schema:accessibilityFeature">tableOfContents</meta>
    <meta property="schema:accessibilityFeature">MathML</meta>
    <meta property="schema:accessibilityFeature">describedMath</meta>
    <meta property="schema:accessibilityAPI">ARIA</meta>
  </metadata>
```

Metadados gerados após conversão da obra:

```
content.opf
1  <package xmlns="http://www.idpf.org/2007/opf" version="3.0" unique-identifier="uuid_id" prefix="calibre: https://calibre-ebook.com">
2    <metadata xmlns:opf="http://www.idpf.org/2007/opf" xmlns:dc="http://purl.org/dc/elements/1.1/" xmlns:dcterms="http://purl.org/dc/terms/"
3      <dc:title id="id">Livro Modelo 4</dc:title>
4      <dc:creator id="id-1">Autor PNLD</dc:creator>
5      <dc:rights>Todos os direitos reservados</dc:rights>
6      <dc:identifier>calibre:30</dc:identifier>
7      <dc:identifier>uuid:bf199d10-853e-4df2-85c2-dafc3db2743a</dc:identifier>
8      <dc:identifier id="uuid_id">uuid:bf199d10-853e-4df2-85c2-dafc3db2743a</dc:identifier>
9      <dc:language>pt</dc:language>
0      <dc:contributor id="id-2">calibre (5.23.0) [https://calibre-ebook.com]</dc:contributor>
1      <dc:date>2021-08-03T03:00:00+00:00</dc:date>
2      <dc:description>Esse é um livro exemplo para teste do Calibre</dc:description>
3      <dc:publisher>Editora PNLD</dc:publisher>
4      <dc:subject>Exemplo</dc:subject>
5      <dc:subject>Calibre</dc:subject>
6      <dc:subject>Conversão</dc:subject>
7      <opf:meta refines="#id" property="title-type">main</opf:meta>
8      <opf:meta refines="#id" property="file-as">Livro Modelo 4</opf:meta>
9      <meta name="cover" content="cover"/>
0      <meta property="calibre:timestamp" scheme="dcterms:W3CDTF">2021-09-01T13:22:18Z</meta>
1      <meta property="dcterms:modified" scheme="dcterms:W3CDTF">2021-09-01T13:22:36Z</meta>
2      <opf:meta refines="#id-2" property="role" scheme="marc:relators">bkp</opf:meta>
3      <opf:meta refines="#id-1" property="role" scheme="marc:relators">aut</opf:meta>
4      <opf:meta refines="#id-1" property="file-as">PNLD, Autor</opf:meta>
5      <opf:meta property="calibre:author_link_map">{"Autor PNLD": ""}</opf:meta>
6    </metadata>
7    <manifest>
```

Apesar de não ter sido identificado nenhum problema nos testes realizados relacionados à falta dos metadados no arquivo OPF referentes a acessibilidade, é importante assegurar que esse

processo ocorra de maneira correta já que alguns leitores e/ou publicadoras validam a acessibilidade da obra por meio desses dados.

5.4.4 Validação do PDF

Para validação do PDF, foi utilizado a aplicação PDF Accessibility Checker (PAC 3), que analisa tecnicamente se todos os elementos de acessibilidade estão presentes no resultado final.

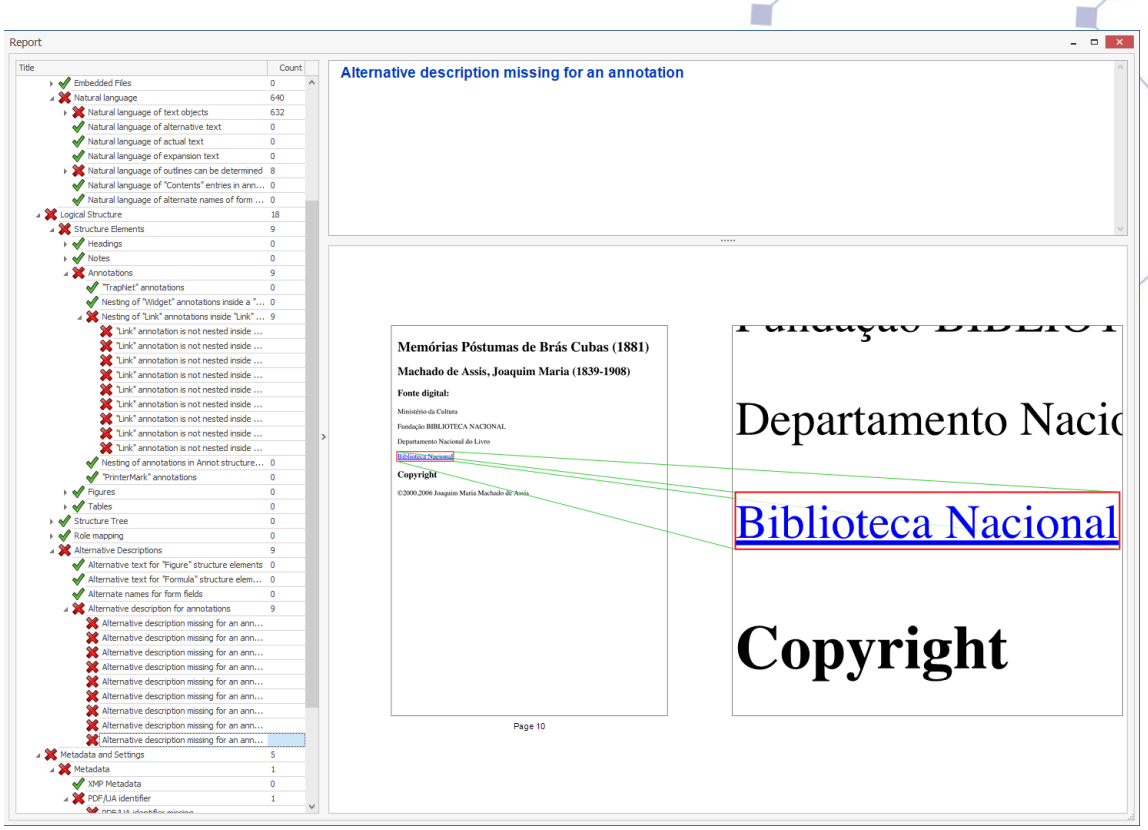
Ao observar os detalhes dos problemas encontrados, alguns se mostram de fácil resolução. Trabalhando apenas no código HTML da obra de origem, podem ser resolvidos. Outros se

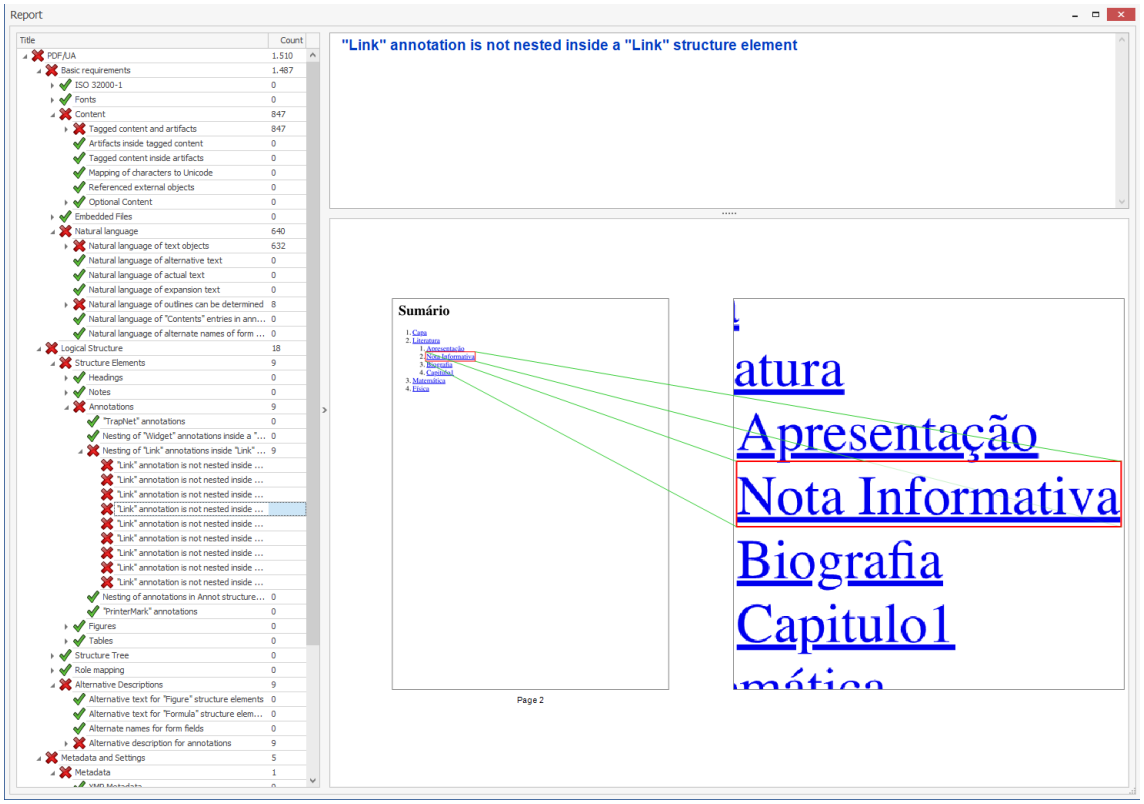
mostram mais complexos, provavelmente devido a forma com que o PDF foi gerado pelo conversor.

Report

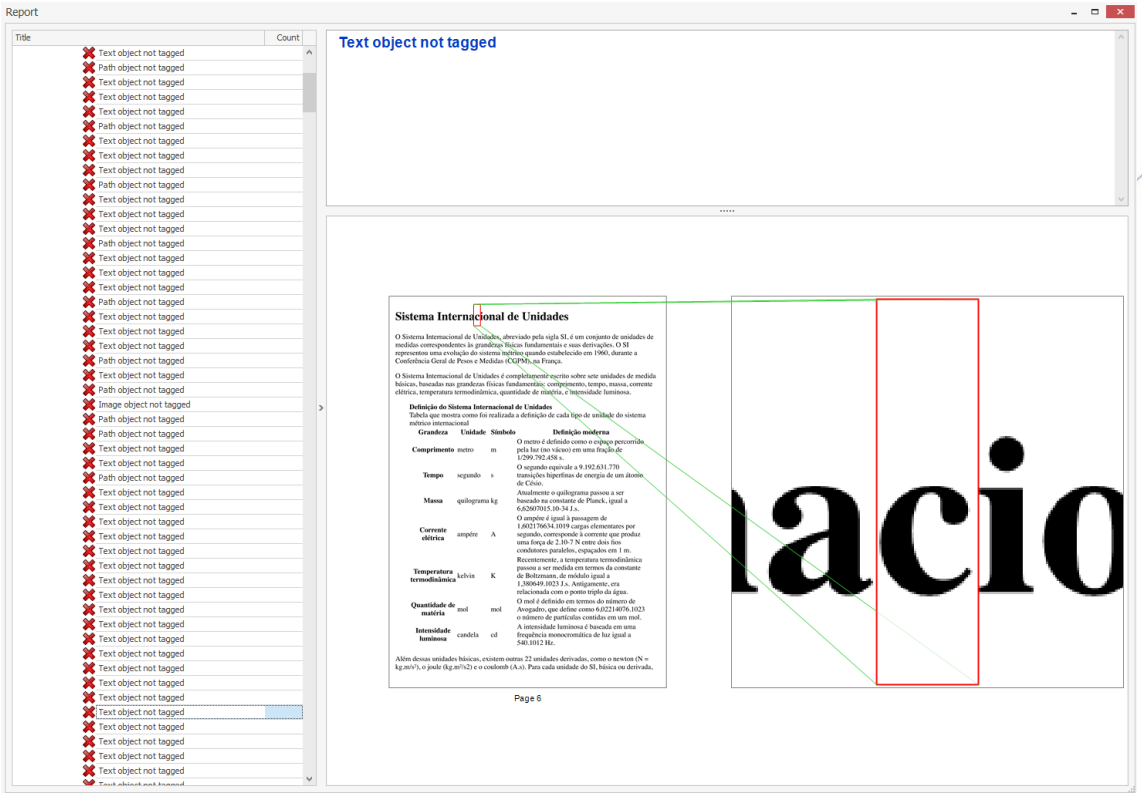
Title	Count
PDF/UA	1.510
Basic requirements	1.487
ISO 32000-1	0
Fonts	0
Content	847
Tagged content and artifacts	847
Artifacts inside tagged content	0
Tagged content inside artifacts	0
Mapping of characters to Unicode	0
Referenced external objects	0
Optional Content	0
Embedded Files	0
Natural language	640
Natural language of text objects	632
Natural language of alternative text	0
Natural language of actual text	0
Natural language of expansion text	0
Natural language of outlines can be determined	8
Natural language of "Contents" entries in anno...	0
Natural language of alternate names of form fi...	0
Logical Structure	18
Structure Elements	9
Headings	0
Notes	0
Annotations	9
Figures	0
Tables	0
Structure Tree	0
Role mapping	0
Alternative Descriptions	9
Alternative text for "Figure" structure elements	0
Alternative text for "Formula" structure elements	0
Alternate names for form fields	0
Alternative description for annotations	9
Metadata and Settings	5
Metadata	1
XMP Metadata	0
PDF/UA identifier	1
PDF/UA identifier missing	
Title in XMP metadata	0
Document settings	4
Display of document title in window title	1
Tag suspects	0
Mark for tagged documents	1
Dynamic XFA form	0
Security settings and document access by assi...	0
Tab order for pages with annotations	2

Nos problemas que podem ser resolvidos via observação no código original, podemos citar a estrutura correta nos *links* e a adição de textos alternativos para navegação, na qual podemos adicionar tags e propriedades no HTML que serão lidas pelo conversor corretamente.


Report
Title
Count
Alternative description missing for an annotation
Memórias Póstumas de Brás Cubas (1881)
Machado de Assis, Joaquim Maria (1839-1908)
Fonte digital:
Copyright
Page 10
Departamento Nacio
Biblioteca Nacional
Copyright


Report
Title
Count
"Link" annotation is not nested inside a "Link" structure element
Sumário
Page 2
atura
Apresentação
Nota Informativa
Biografia
Capitulo1

Porém, outros problemas se apresentam um pouco mais complexos como a tag em objetos.



Ao converter o material para PDF, o conteúdo é separado adequadamente em blocos, mas nenhuma tag informativa, descritiva ou anotação é gerada no conteúdo final e nos testes realizados não foi encontrada uma forma de fazer essa conversão. Não foi possível apurar qual o método/ferramenta o Calibre utiliza para realizar essa conversão, contudo, é provável que, para a conversão de um PDF acessível, irá ser preciso ajustes ou elaboração de um *plugin* próprio para garantir a qualidade do material.

5.4.5 Conclusão: Livro Modelo 4

Ao observar os resultados após a conversão do material, fica evidente que a conversão de um livro acessível utilizando o calibre é possível e, de um modo geral, traz bons resultados. Porém, é importante ressaltar que, para que se tenha uma conversão perfeita de uma obra acessível, é preciso tomar alguns cuidados na elaboração da obra e principalmente trabalhar em um modelo de conversão para PDF mais efetivo.

Como o Calibre é uma ferramenta *open source* e que permite a adição de *plugins* para realização das tarefas de conversão, no futuro, pode-se trabalhar em um conversor próprio para PDF que detenha todas as características necessárias para execução dessa tarefa, podendo inclusive se basear em ferramentas já existem como https://github.com/danfickle/openhtmltopdf/wiki/PDF-Accessibility-(PDF-UA,-WCAG,-

Section-508)-Support, que faz a conversão de HTML para PDF seguindo os padrões de acessibilidade.

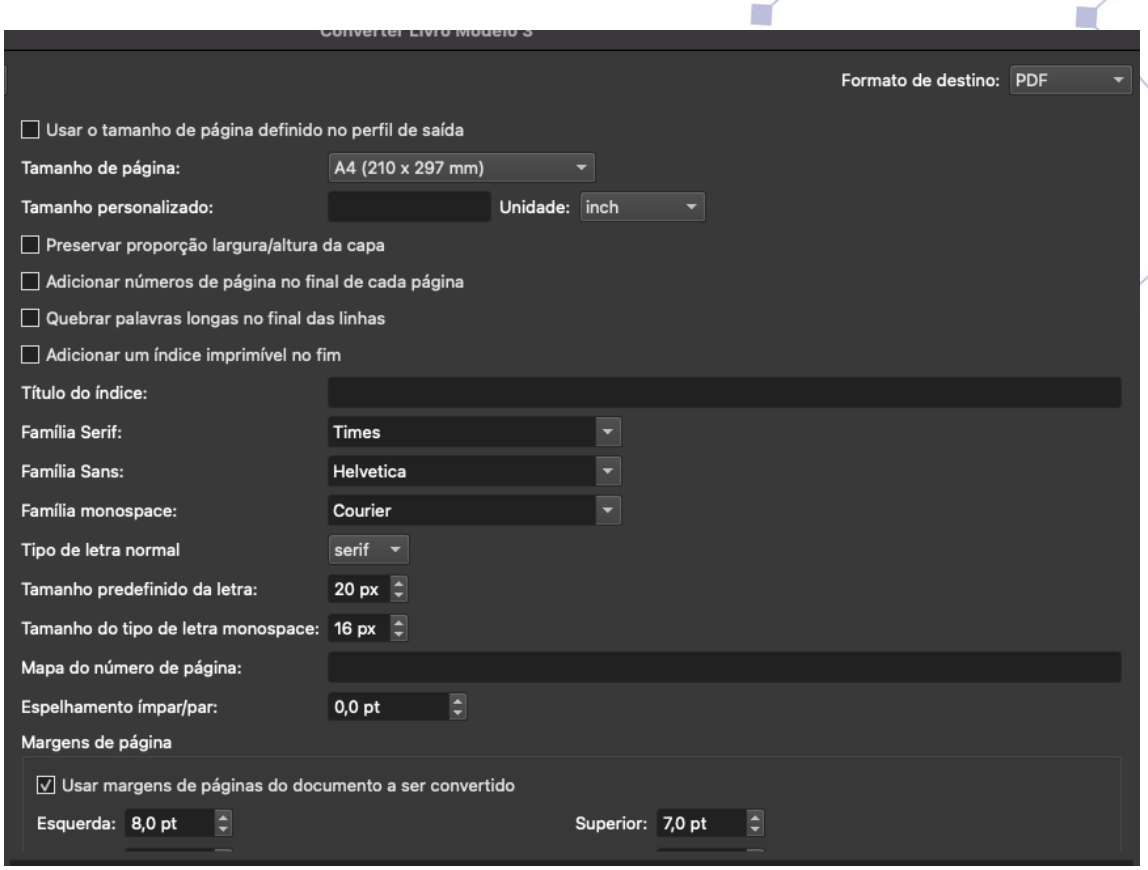
6. Conclusão

A partir dessa prova de conceito, pode-se concluir que o Calibre será a ferramenta mais adequada para garantir que a obra possa ser disponibilizada nos mais diversos formatos para todos os usuários. Por se tratar de uma ferramenta *open source* e que permite a instalação de *plugins* de terceiros para incrementar suas opções, será crucial para expansão de opções futuras para as obras disponibilizadas pela plataforma do PNLD, garantindo que seja possível evoluir o processo de acordo com as necessidades do PNLD.

Importante salientar que a ferramenta conta com uma comunidade ativa trabalhando em processos de melhoria, correções de *bugs* e novas funcionalidades como pode ser observado no repositório oficial da ferramenta: https://github.com/kovidgoyal/calibre.

Com base nos testes apresentados, é importante destacar uma necessidade que poderá ser apresentada futuramente para melhoria da plataforma PNLD. Para garantir que cada editora tenha sua obra convertida conforme os padrões testados no momento do desenvolvimento do projeto, é importante que, no futuro, a plataforma permita que todas as opções de conversão, tanto de entrada como de saída, disponíveis na ferramenta de conversão de linha de comando seja parametrizada na própria plataforma.

Essa parametrização deve seguir a documentação oficial do Calibre, disponível em https://manual.calibre-ebook.com/generated/en/ebook-convert.html, para o formato de entrada em HTML e todos os formatos de saída permitidos pelo sistema. Assim sendo, as editoras podem configurar todos os parâmetros necessários para que sua obra seja convertida da forma correta, assim como o Calibre permite fazer algumas dessas configurações por meio de sua interface gráfica.

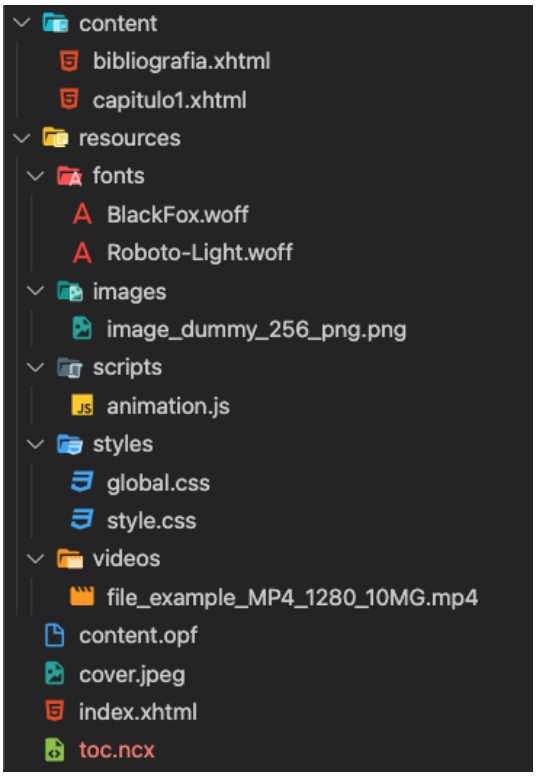


Outro ponto importante para o projeto está na padronização das obras para garantir que todas as editoras tenham acesso a um modelo estável e testado para criação do projeto. Essa padronização visa a garantir que a plataforma possa fazer as tratativas necessárias, seja de conversão da obra para outros formatos ou análises mais robustas dos arquivos de maneira automatizada.

Para isso, é importante que todas as obras sejam desenvolvidas utilizando XHTML5 e CSS 3 como padrão, apesar dessas tecnologias não garantirem a maior interoperabilidade entre os formatos, já que o EPUB2 não aceita as tags em HTML5. No entanto, será possível avançar a tecnologia com os leitores mais recentes e trabalhar com os modelos de acessibilidade e semântica de conteúdo que são a base dessas tecnologias.

Importante salientar a informação trazida pelo Calibre da obrigação de utilização do XHTML ao invés do HTML para um melhor resultado na conversão das obras. Apesar de não observarmos nenhuma diferença nos testes executados, o XHTML garante que todo código escrito seja um XML válido e, nesse caso, é passível de validação e evita erros futuros.

Na padronização do projeto, é aconselhável a utilização da seguinte estrutura:



Sendo:

- toc.ncx - Definição do sistema de navegação da obra (Obrigatório).
- cover.jpeg - Capa da obra (Obrigatório).
- content.opf - Definição dos metadados da obra (Obrigatório).
- index.xhtml - Página inicial da obra (Obrigatório).
- content - Pasta na qual serão armazenados a diagramação da obra, separada por capítulos (Obrigatório).
- resources - Pasta em que serão armazenados todos os recursos utilizados na obra (Opcional).

É importante que todos os nomes dos arquivos e pastas da estrutura da obra sejam descritos no idioma inglês, garantindo assim que seja possível identificar todos os elementos de forma automatizada e universal. Manter os principais nomes em inglês traz também o benefício sobre as ferramentas de desenvolvimento que, de um modo geral, interpretam esses identificadores e já fazem uma apresentação visual diferenciada, como os ícones apresentados em cada pasta, de maneira automática.

De um modo geral, essa prova de conceito demonstra a importância e robustez do calibre para realização da conversão das obras, mas com algumas deficiências que precisarão ser tratadas no futuro, principalmente com relação a conversão de obras para PDF.

7. Referências

Calibre. Disponível em: <https://calibre-ebook.com/>. Acesso em: 02 ago. 2021.

Calibre **CLI**. Disponível em: <https://manual.calibre-ebook.com/generated/en/cli-index.html>. Acesso em: 02 ago. 2021.

Open Packaging Format - **OPF**: Disponível em <http://idpf.org/epub/20/spec/OPF_2.0.1_draft.htm#Section2.2>. Acesso em: 02 ago. 2021.

Documentação da ferramenta **ebook_convert**: Disponível em <https://manual.calibre-ebook.com/generated/en/ebook-convert.html>. Acesso em: 02 ago. 2021.

Web Content Accessibility Guidelines **WCAG**. Disponível em: <https://www.w3.org/WAI/standards-guidelines/wcag/>. Acesso em: 02 ago. 2021.

International Digital Publishing Forum - **IDPF**. Disponível em: <http://idpf.org/>. Acesso em: 02 ago. 2021.

EPUB Validator (Beta). Disponível em: <http://validator.idpf.org/>. Acesso em: 02 ago. 2021.

Amazon **Kindle** Publishing Guidelines. Disponível em: <https://kindlegen.s3.amazonaws.com/AmazonKindlePublishingGuidelines.pdf>. Acesso em 05 ago. 2021.

Aria. Disponível em <https://developer.mozilla.org/pt-BR/docs/Web/Accessibility/ARIA>. Acesso em 01 set. 2021.

8. Anexos

8.1 Exemplo arquivo toc.ncx.

Exemplo de arquivo no formato ncx - Navigation Control file for XML para descrição da tabela de conteúdo da obra.

```
<?xml version="1.0"?>
<!DOCTYPE ncx PUBLIC "-//NISO//DTD ncx 2005-1//EN"
"http://www.daisy.org/z3986/2005/ncx-2005-1.dtd">
<ncx xmlns="http://www.daisy.org/z3986/2005/ncx/" version="2005-1">
 <head></head>
 <docTitle>
   <text>Livro Modelo 4</text>
 </docTitle>
 <navMap>
   <!-- 1 -->
   <navPoint class="chapter" id="toc" playOrder="1">
     <navLabel>
       <text>Sumário</text>
     </navLabel>
     <content src="index.html#toc" />
   </navPoint>
   <!-- 2 -->
   <navPoint class="chapter" id="capa" playOrder="2">
     <navLabel>
       <text>capa</text>
     </navLabel>
     <content src="content/capa.xhtml" />
   </navPoint>
   <!-- 3 -->
   <navPoint class="chapter" id="literatura" playOrder="3">
     <navLabel>
       <text>Literatura</text>
     </navLabel>
     <content src="content/literatura.xhtml" />
   </navPoint>
   <!-- 4 -->
   <navPoint class="section" id="literatura_apr" playOrder="4">
     <navLabel>
       <text>Apresentação</text>
     </navLabel>
     <content src="content/literatura.xhtml#presentation" />
   </navPoint>
   <!-- 5 -->
   <navPoint class="section" id="literatura_note" playOrder="5">
     <navLabel>
       <text>Nota informativa</text>
     </navLabel>
     <content src="content/literatura.xhtml#note" />
   </navPoint>
   <!-- 6 -->
   <navPoint class="section" id="literatura_bio" playOrder="6">
     <navLabel>
       <text>Biografia</text>
     </navLabel>
     <content src="content/literatura.xhtml#biography" />
   </navPoint>
   <!-- 7 -->
   <navPoint class="section" id="literatura_cap1" playOrder="7">
     <navLabel>
```

```
        <text>Capitulo 1</text>
      </navLabel>
      <content src="content/literatura.xhtml#capitulo1" />
    </navPoint>
    <!-- 8 -->
    <navPoint class="chapter" id="matematica" playOrder="8">
      <navLabel>
        <text>Matemática</text>
      </navLabel>
      <content src="content/matematica.xhtml" />
    </navPoint>
    <!-- 9 -->
    <navPoint class="chapter" id="fisica" playOrder="9">
      <navLabel>
        <text>Física</text>
      </navLabel>
      <content src="content/fisica.xhtml" />
    </navPoint>
  </navMap>
</ncx>
```

8.2 Exemplo arquivo content.opf

Exemplo de arquivo no formato Open Packaging Format para implementação dos metadados da obra.

```
<?xml version="1.0" encoding="utf-8" ?>
<package     version="2.0"     xmlns="http://www.idpf.org/2007/opf"     unique-
identifier="BookId">
 <metadata                          xmlns:dc="http://purl.org/dc/elements/1.1/"
xmlns:opf="http://www.idpf.org/2007/opf">
   <dc:title>Livro Modelo 4</dc:title>
   <dc:language>pt-br</dc:language>

   <dc:creator opf:role="aut" opf:file-as="PNLD, Autor">Autor PNLD</dc:creator>
   <dc:publisher>Editora PNLD</dc:publisher>
   <dc:date>2021-08-03</dc:date>

   <dc:description>Esse    é    um    livro    exemplo    para    teste    do
Calibre</dc:description>

   <dc:subject>Exemplo</dc:subject>
   <dc:subject>Calibre</dc:subject>
   <dc:subject>Conversão</dc:subject>

   <dc:rights>Todos os direitos reservados</dc:rights>

   <meta name="cover" content="cover.jpeg" />
   <x-metadata>
     <EmbeddedCover>cover.jpeg</EmbeddedCover>
   </x-metadata>

   <meta property="schema:accessibilityFeature">structuralNavigation</meta>
   <meta property="schema:accessibilityFeature">tableOfContents</meta>
   <meta property="schema:accessibilityFeature">MathML</meta>
   <meta property="schema:accessibilityFeature">describedMath</meta>
   <meta property="schema:accessibilityAPI">ARIA</meta>
 </metadata>

 <manifest>
   <item id="content" href="index.xhtml" media-type="application/xhtml+xml" />
   <item        id="content_capa"        href="content/capa.xhtml"        media-
type="application/xhtml+xml" />
   <item    id="content_literatura"    href="content/literatura.xhtml"    media-
type="application/xhtml+xml" />
   <item    id="content_matematica"    href="content/matematica.xhtml"    media-
type="application/xhtml+xml" />
   <item      id="content_fisica"      href="content/fisica.xhtml"      media-
type="application/xhtml+xml" />
   <item id="ncx" href="toc.ncx" media-type="application/x-dtbncx+xml" />
   <item    id="cover"    properties="cover-image"    href="cover.jpeg"    media-
type="image/jpeg" />
 </manifest>

 <spine toc="ncx">
   <itemref idref="content" />
 </spine>

 <guide>
   <reference href="index.xhtml#toc" type="toc" title="Table of Contents" />
   <reference href="content/literatura.xhtml" type="text" title="Literatura" />
   <reference href="content/matematica.xhtml" type="text" title="Matemática" />
   <reference href="content/fisica.xhtml" type="text" title="Física" />
 </guide>
```

```
</package>
```

8.3 Exemplo arquivo index.xhtml

Exemplo de arquivo html para geração do sumário e sistema de navegação base da obra. Nesse exemplo foram substituídas as tags *<nav>* e *<main>* por uma * * para garantir a aprovação do formato EPUB 2.

```
<?xml version='1.0' encoding='utf-8'?>
<!DOCTYPE html>
<html lang="pt-br" xml:lang="pt-br" xmlns="http://www.w3.org/1999/xhtml">
 <head>
   <meta charset="UTF-8" />
   <title>Nome do Livro</title>
   <meta
     name="description"
     content="Esta é uma obra modelo para demonstrar o conteúdo
             de um HTML5 formatado de acordo com o modelo do do PNLD."
   />
   <meta
     name="generator"
     content="Ferramenta usada para a geração do documento."
   />
   <meta name="author" content="Teste Calibre" />

   <meta name="robots" content="noindex, nofollow" />
 </head>

 <body lang="pt-BR" xml:lang="pt-BR">
   <div itemscope="" itemtype="https://schema.org/Book">
     <meta itemprop="accessibilityFeature" content="largePrint/CSSEnabled" />
     <meta itemprop="accessibilityFeature" content="highContrast/CSSEnabled" />
     <meta itemprop="accessibilityFeature" content="resizeText/CSSEnabled" />
     <meta itemprop="accessibilityFeature" content="displayTransformability" />
     <meta itemprop="accessibilityFeature" content="longDescription" />
     <meta itemprop="accessibilityFeature" content="alternativeText" />
     <meta itemprop="accessibilityControl" content="fullKeyboardControl" />

     <meta itemprop="name" content="Livro Modelo para o PNLD 2023" />
     <meta
       itemprop="description"
       content="Este é um livro modelo para o PNLD 2023"
     />
     <meta itemprop="isbn" content="9780030426879" />

     <meta
       itemprop="publisher"
       itemtype="https://schema.org/Organization"
       itemscope=""
       content="Grupo de Trabalho para o PNLD23"
     />

     <meta itemprop="copyrightYear" content="Livro Modelo para o PNLD 2023" />
     <meta
       itemprop="copyrightHolder"
       content="Livro Modelo para o PNLD 2023"
     />
     <meta itemprop="inLanguage" content="Livro Modelo para o PNLD 2023" />

     <meta
       itemprop="author"
       itemtype="https://schema.org/Organization"
       itemscope=""
       content="Grupo de Trabalho para o PNLD23"
```

```
      />
    </div>
    <nav role="doc-toc" id="toc">
      <h1>Sumário</h1>
      <ol>
        <li><a href="content/capa.xhtml">Capa</a></li>
        <li>
          <a href="content/literatura.xhtml">Literatura</a>
          <ol>
            <li>
              <a href="content/literatura.xhtml#presentation">Apresentação</a>
            </li>
            <li>
              <a href="content/literatura.xhtml#note">Nota Informativa</a>
            </li>
            <li><a href="content/literatura.xhtml#biography">Biografia</a></li>
            <li><a href="content/literatura.xhtml#capitulo1">Capitulo1</a></li>
          </ol>
        </li>
        <li>
          <a href="content/matematica.xhtml">Matemática</a>
        </li>
        <li>
          <a href="content/fisica.xhtml">Física</a>
        </li>
      </ol>
    </nav>
  </body>
</html>
```

8.4 Exemplo arquivo de conteúdo content.xhtml

Exemplo de arquivo html para geração do conteúdo da obra.

```
<?xml version='1.0' encoding='utf-8'?>
<!DOCTYPE html>
<html lang="pt-BR" xml:lang="pt-BR" xmlns="http://www.w3.org/1999/xhtml">
 <head>
   <meta charset="UTF-8" />
   <title>Memórias póstumas de brás cubas</title>
   <meta name="robots" content="noindex, nofollow" />
 </head>

 <body lang="pt-BR" xml:lang="pt-BR">
   <div itemscope="" itemtype="https://schema.org/Book">
     <meta itemprop="accessibilityFeature" content="largePrint/CSSEnabled" />
     <meta itemprop="accessibilityFeature" content="highContrast/CSSEnabled" />
     <meta itemprop="accessibilityFeature" content="resizeText/CSSEnabled" />
     <meta itemprop="accessibilityFeature" content="displayTransformability" />
     <meta itemprop="accessibilityFeature" content="longDescription" />
     <meta itemprop="accessibilityFeature" content="alternativeText" />
     <meta itemprop="accessibilityControl" content="fullKeyboardControl" />
   </div>

   <main>
     <!-- Conteudo -->
   </main>
 </body>
</html>
```