

Conversão de Markdown para DOCX e PDF

Este arquivo documenta a conversão dos arquivos Markdown do artigo científico para os formatos DOCX e PDF.

Status da Conversão - Atualizado em 2026-01-02

Fase 4 (fase4_secoes)

Todos os 12 arquivos Markdown desta pasta foram convertidos com sucesso:

Arquivo Markdown	DOCX	PDF	Tamanho PDF
agradecimentos_referencias.md			69 KB
conclusao_completa.md			70 KB
contraprova_casos_limite.md			70 KB
discussao_completa.md			135 KB
introducao_completa.md			87 KB
metodologia_completa.md			124 KB
prova_teorema.md			73 KB
resultados_completo.md			110 KB
resumo_abstract.md			43 KB
revisao_literatura_completa.md			108 KB
secao_didatica_leigos.md			79 KB
teorema_beneficio_condicionado.md			91 KB

Total: 12 arquivos MD → 12 DOCX + 12 PDF = 24 arquivos de saída

Status Completo - Todas as Pastas

A conversão foi expandida para incluir **todos** os arquivos Markdown na pasta artigo_cientifico:

Pasta	Arquivos MD	DOCX	PDF	Taxa de Sucesso
fase1_analise	2	2	2	100%
fase2_bibliografia	2	2	2	100%
fase3_estrutura	2	2	2	100%
fase4_secoes	13	13	13	100%
fase5_suplementar	9	9	7	78% PDF
fase6_consolidacao	9	9	9	100%
latex_template	1	1	1	100%
raiz	8	8	8	100%
TOTAL	46	** 46**	** 44**	95.7%

Arquivos com Problemas na Conversão PDF

Apenas 2 arquivos tiveram problemas na conversão para PDF devido a padrões LaTeX complexos:

1. **fase5_suplementar/apendice_g_validacao_estatistica.md**
 - DOCX criado com sucesso
 - PDF com erro LaTeX (comando \text@ incompleto)
 - Solução: Revisar manualmente os comandos \text{} no arquivo
2. **fase5_suplementar/apendice_i_lista_simbolos.md**
 - DOCX criado com sucesso
 - PDF com erro LaTeX (notação de norma \l em tabelas)
 - Solução: Simplificar notação matemática nas tabelas

Scripts de Conversão

Script Original (apenas fase4_secoes)

```
python3 tools/convert_fase4_to_docx_pdf.py
```

Novo Script (todas as pastas recursivamente)

```
# Converter TODOS os arquivos MD em artigo_cientifico
```

```
python3 tools/convert_all_artigo_mds.py
```

```
# Ou especificar uma pasta específica
```

```
python3 tools/convert_all_artigo_mds.py artigo_cientifico/fase5_suplementar
```

Como Usar os Scripts de Conversão

Pré-requisitos

Os scripts requerem: - Python 3.x - Pandoc (instalado via `apt-get install pandoc`) - LaTeX (texlive-xetex) para geração de PDF

Instalação das Dependências

```
# Instalar pandoc
```

```
sudo apt-get update
```

```
sudo apt-get install -y pandoc
```

```
# Instalar LaTeX para PDFs
```

```
sudo apt-get install -y texlive-xetex texlive-fonts-recommended texlive-plain-generic
```

Execução

Converter APENAS fase4_secoes (script original):

```
python3 tools/convert_fase4_to_docx_pdf.py
```

```
# Ou especificar outro diretório
```

```
python3 tools/convert_fase4_to_docx_pdf.py /caminho/para/diretorio
```

Converter TODAS as pastas recursivamente (novo script):

```
# Converter todos os arquivos .md da pasta artigo_cientifico e subpastas
```

```
python3 tools/convert_all_artigo_mds.py
```

```
# Ou especificar outro diretório
```

```
python3 tools/convert_all_artigo_mds.py /caminho/para/diretorio
```

Características dos Scripts

Script Original (convert_fase4_to_docx_pdf.py): - Converte apenas arquivos em um diretório específico (não recursivo) - Padrão: `artigo_cientifico/fase4_secoes`

Novo Script (convert_all_artigo_mds.py): - Busca recursiva em todas as subpastas - Agrupa resultados por diretório para melhor visualização - Relatório detalhado com estatísticas por pasta - Padrão: `artigo_cientifico` (todas as subpastas)

Ambos os scripts: 1. **Encontram automaticamente** todos os arquivos `.md` 2. **Convertem para DOCX** usando pandoc com suporte a: - Tabelas - Blocos de código - Formatação Markdown completa 3. **Convertem para PDF** usando pandoc + XeLaTeX com: - Suporte completo a Unicode

- Renderização de equações LaTeX (formato \dots e \dots) - Fontes DejaVu para melhor compatibilidade - Margens de 2cm 4. **Sanitizam o conteúdo** automaticamente: - Removem emojis problemáticos para LaTeX - Convertem checkboxes markdown (- [DONE] → - [DONE]) - Corrigem padrões de fórmulas matemáticas que causam erros 5. **Geram relatório** detalhado com estatísticas de sucesso/falha

Formato dos Arquivos

DOCX

- Arquivos editáveis no Microsoft Word, LibreOffice, Google Docs
- Mantém formatação básica (títulos, listas, tabelas, negrito, itálico)
- Tamanho típico: 15-35 KB

PDF

- Formato final para leitura e distribuição
- Renderização completa de equações matemáticas via LaTeX
- Formatação profissional com fontes serifadas
- Tamanho típico: 40-140 KB dependendo do conteúdo matemático

Observações Técnicas

Equações Matemáticas

Os arquivos suportam LaTeX inline e display:

Equação inline: $E = mc^2$

Equação em bloco:

```
$$
\hat{H} = \sum_i \sigma_i^z
$$
```

Tratamento de Caracteres Especiais

O script remove automaticamente: - Emojis (, , , etc.) que causam problemas no LaTeX - Checkboxes markdown são convertidos para texto

Engine de PDF

O script usa xelatex como engine de PDF porque: - Suporte nativo a Unicode - Melhor renderização de caracteres especiais - Compatibilidade com fontes TrueType/OpenType

Manutenção

Para reconverter todos os arquivos (por exemplo, após atualizar o Markdown):

Apenas fase4_secoes:

```
# Remover conversões antigas
```

```
rm artigo_cientifico/fase4_secoes/*.docx artigo_cientifico/fase4_secoes/*.pdf
```

```
# Reconverter
```

```
python3 tools/convert_fase4_to_docx_pdf.py
```

Todas as pastas:

```
# Remover todas as conversões antigas
find artigo_cientifico -type f \( -name "*.docx" -o -name "*.pdf" \) -delete

# Reconverter tudo
python3 tools/convert_all_artigo_mds.py
```

Solução de Problemas

Erro: “pandoc: command not found”

```
sudo apt-get install pandoc
```

Erro: “xelatex not found”

```
sudo apt-get install texlive-xetex texlive-fonts-recommended
```

PDF não gera corretamente

- Verifique se há caracteres especiais não suportados no Markdown
- O script já trata os casos mais comuns, mas novos padrões podem precisar de ajustes
- Verifique os logs de erro do pandoc para detalhes

Estrutura de Arquivos

```
artigo_cientifico/fase4_secoes/
├── *.md                # Arquivos Markdown originais
├── *.docx              # Conversões DOCX
├── *.pdf               # Conversões PDF
└── README_CONVERSION.md # Esta documentação
```

Contato

Para problemas ou sugestões relacionadas à conversão, consulte os scripts em: - tools/convert_fase4_to_docx (script original, apenas fase4_secoes) - tools/convert_all_artigo_mds.py (novo script, conversão recursiva completa)

Última atualização: 02 de janeiro de 2026

Versão do script: 2.0 (adicionado suporte recursivo)

Status: 44/46 arquivos convertidos com sucesso (95.7%)