

Curadoria de Referências

Um guia prático para garantir qualidade, confiabilidade e padronização das suas referências bibliográficas

Eduardo Ogasawara
eduardo.ogasawara@cefet-rj.br
<https://eic.cefet-rj.br/~eogasawara>

Objetivo da Curadoria

A curadoria de referências é um processo essencial que visa garantir a qualidade e confiabilidade das fontes citadas em trabalhos acadêmicos. Este procedimento assegura a correta identificação das referências em bases de dados reconhecidas internacionalmente e promove a padronização e completude dos metadados bibliográficos.



Qualidade e Confiabilidade

Garante que todas as fontes citadas sejam verificadas e validadas em bases reconhecidas



Identificação Correta

Assegura que as referências estejam devidamente indexadas e identificáveis



Padronização de Metadados

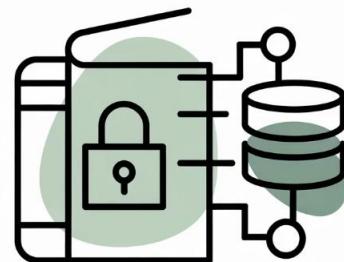
Uniformiza e completa as informações bibliográficas essenciais



Além de ser fundamental para garantir a reproduzibilidade da pesquisa, a curadoria adequada facilita futuras buscas bibliográficas e aumenta significativamente a credibilidade do seu trabalho científico.

Por que fazer curadoria de referências?

A curadoria de referências é uma etapa crítica no processo de produção científica que garante a integridade e a qualidade das citações utilizadas em sua pesquisa.



Verificações Essenciais

- Verificar se há acesso ao artigo completo em texto integral
- Conferir se os metadados (autor, título, periódico, ano, volume, páginas) estão corretos e completos
- Certificar que os artigos estão devidamente indexados em bases confiáveis e reconhecidas pela comunidade científica

1

Scopus

Base multidisciplinar da Elsevier,
amplamente reconhecida pela CAPES

2

IEEE Xplore

Biblioteca digital especializada em
engenharia e tecnologia

3

ACM Digital Library

Repositório essencial para ciência da
computação e áreas correlatas

Instalação do Zotero (Passo a Passo)

O Zotero é uma ferramenta gratuita e de código aberto para gerenciamento de referências bibliográficas. Sua instalação e configuração adequadas são fundamentais para o processo de curadoria.

01

Download do Aplicativo

Acesse zotero.org e baixe o aplicativo para seu sistema operacional (Windows, Mac ou Linux)

02

Extensão do Navegador

Instale a extensão Zotero Connector para Firefox ou Chrome para captura automática de referências

03

BibTeX.js Customizado

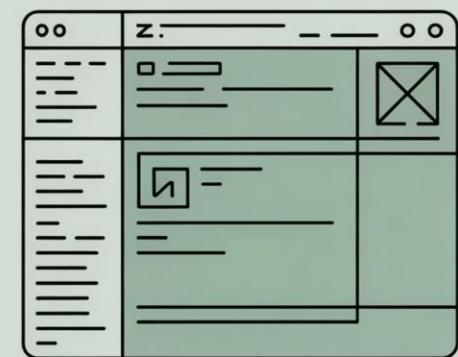
Baixe o arquivo BibTeX.js do repositório disponível em <https://github.com/eogasawara/mylibrary>

04

Instalação do Tradutor

Copie o arquivo BibTeX.js para a pasta de tradução do seu Zotero (Ex.: /Zotero/translators)

- Importante:** Refaça o processo de instalação do tradutor customizado a cada atualização do Zotero para garantir compatibilidade.



Configuração e Sincronização do Zotero

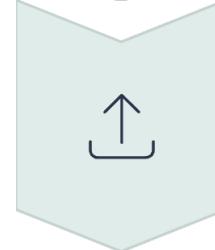


A sincronização na nuvem permite acessar sua biblioteca de referências de qualquer dispositivo e colaborar com colegas de pesquisa de forma eficiente.



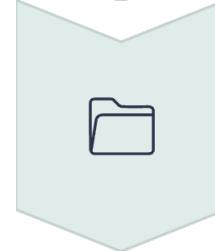
Criação de Conta

Crie uma conta gratuita em <https://www.zotero.org> para habilitar a sincronização



Ativação da Sincronização

Configure a sincronização automática nas preferências do aplicativo para backup na nuvem



Organização Estratégica

Organize sua biblioteca criando uma pasta específica por artigo e uma pasta para cada TCC, dissertação ou tese

- ☐ **Atenção especial para alunos do DAL:** Os alunos do Data Analytics Lab devem entrar no grupo compartilhado disponível em <https://www.zotero.org/groups/2064360/data-analytics-lab/library>

Configuração do LaTeX

A organização adequada dos arquivos LaTeX é essencial para um fluxo de trabalho eficiente. A estrutura pode variar de simples a complexa, dependendo do tamanho do seu projeto.

Estrutura Básica

Para artigos simples ou projetos pequenos:

```
/paper/|— main.tex|— references.bib
```

Esta estrutura é ideal para artigos curtos onde todo o conteúdo pode ser mantido em um único arquivo principal.

Estrutura Avançada

Para documentos maiores como dissertações ou teses:

```
/paper/|— principal.tex|— parte1.tex|— parte2.tex|— ...|—  
parten.tex|— references.bib
```

Esta abordagem modular facilita a organização e manutenção de documentos extensos.

Preparação do Ambiente

Copie todos os arquivos .tex e .bib para uma pasta de trabalho local (Ex.: /paper) onde você realizará o processo de curadoria

Primeiro uso do Zotero com referências

O primeiro passo no processo de curadoria envolve importar as referências existentes, identificar duplicatas e preparar os dados para consulta nas bases indexadoras.



Importação de Referências

Importe todas as referências bibliográficas do seu artigo para o Zotero através do arquivo .bib ou da extensão do navegador



Verificação de Duplicatas

Utilize a função de detecção automática de duplicatas do Zotero para identificar entradas repetidas na sua biblioteca



Seleção de Metadados

Escolha os metadados mais completos e precisos, frequentemente presentes em versões anteriores já curadas ou importadas de bases confiáveis



Exportação para Consulta

Exporte o arquivo .bib atualizado que será utilizado para gerar a string de consulta ao Scopus

(a) Referências importadas

A screenshot of the Zotero interface showing a library titled 'references'. The left sidebar shows various folder categories like 'BOOKS', 'course', 'dal', 'papers', etc. The main pane displays a list of imported references with columns for Title, Creator, Year, Publication, Rights, Date, and Pages. Several entries are visible, such as 'A bibliometric rev...' by Jalal... from 2021, and 'A Bitcoin ETF Is H...' by Wurs... from 2021.

(b) Verificando duplicatas

A screenshot of the Zotero interface showing the 'Duplicate items' screen. It lists several entries from the 'papers' folder, including 'A bibliometric rev...' by Jalal... from 2021 and 'A Bitcoin ETF Is H...' by Wurs... from 2021. A tooltip indicates 'Merge 2 items' for the first entry. On the right side, there's a panel for selecting fields to keep from other versions of the item, and details about the selected item like title, author, and publication date.

Consulta de curadoria com R

A consulta automatizada através do script R desenvolvido especificamente para curadoria facilita a busca sistemática nas bases de dados. O script gera strings de consulta otimizadas para o Scopus.

Carregar Script

Abra o RStudio e execute o comando para carregar as funções de curadoria

Consulta por DOI

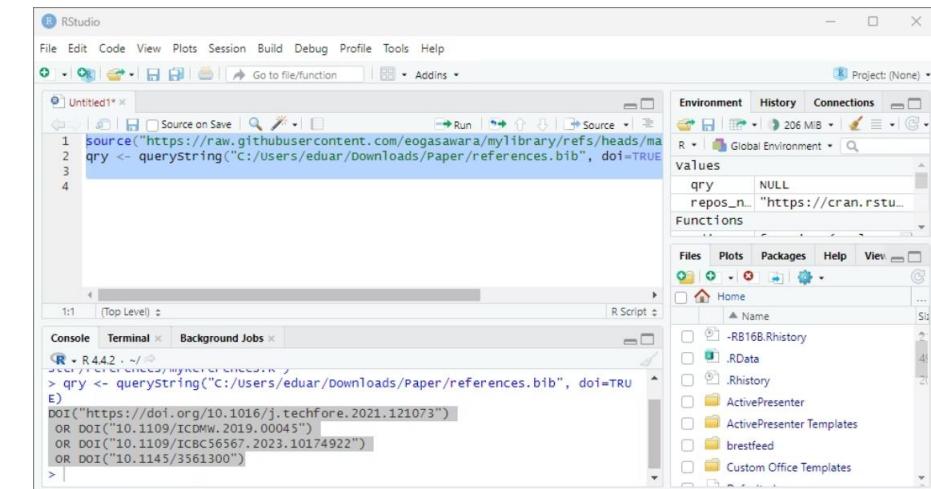
Gere string de busca usando os DOIs disponíveis no arquivo .bib

Consulta por Título

Para referências sem DOI, gere string usando os títulos dos artigos

Comandos Essenciais

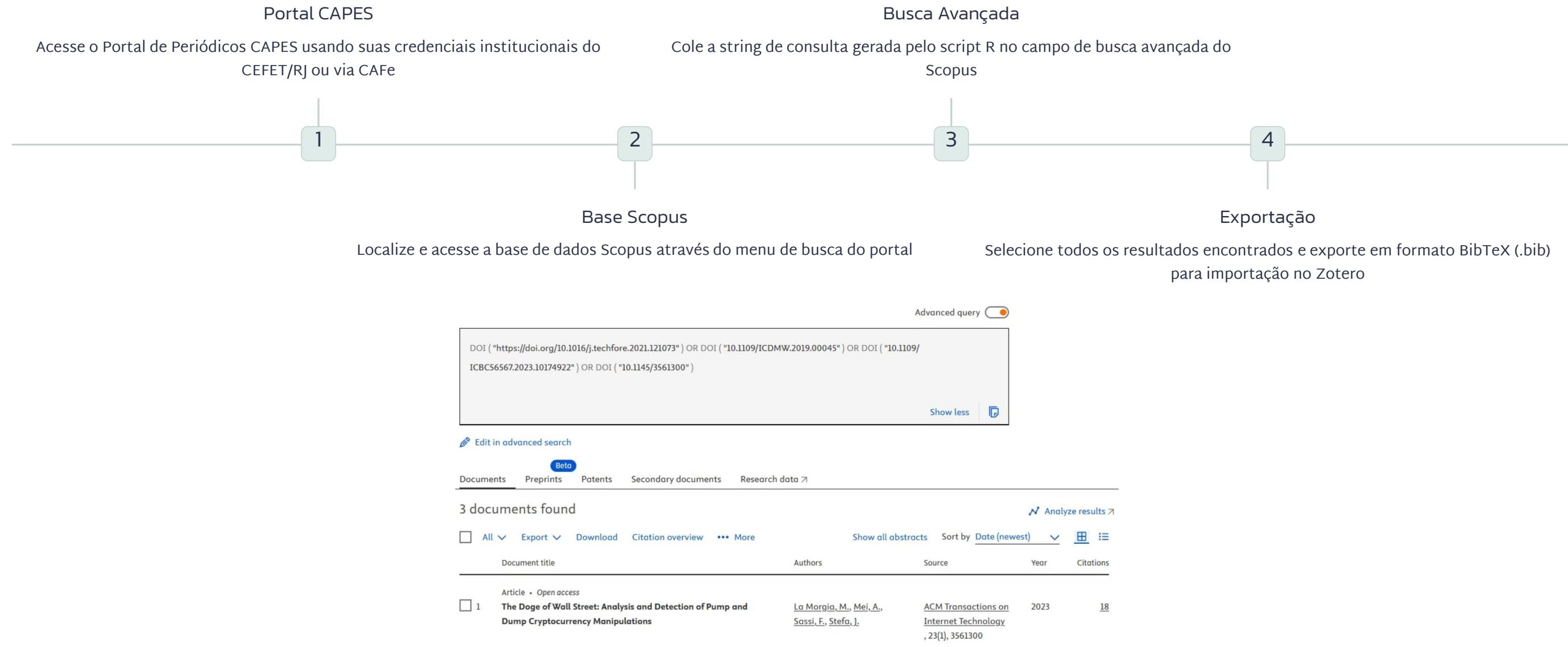
```
# Carregar o script de
curadoriasource("https://raw.githubusercontent.com/eogasawara/mylibrary/refs/he
ads/master/references/myReferences.R")# Consultar artigos pelos
DOIsqueryString("/Paper/references.bib", doi=TRUE)# Consultar artigos pelos
títulos (quando não há DOI)queryString("/Paper/references.bib", doi=FALSE)
```



A saída do script fornece uma string formatada que pode ser copiada diretamente para o campo de busca do Scopus, economizando tempo e reduzindo erros manuais.

Acesso ao Scopus e exportação

O Scopus é uma das principais bases de dados de citações e resumos de literatura científica. O acesso através do Portal CAPES garante privilégios institucionais completos.



- Dica importante:** Salve o arquivo exportado com um nome descritivo como "scopus.bib" para facilitar a identificação durante o processo de mesclagem.

Mesclagem e refinamento de metadados

A mesclagem dos metadados obtidos do Scopus com sua biblioteca existente é uma etapa crucial que enriquece e padroniza suas referências bibliográficas.



Importação

Importe o arquivo scopus.bib recém-exportado para sua biblioteca Zotero

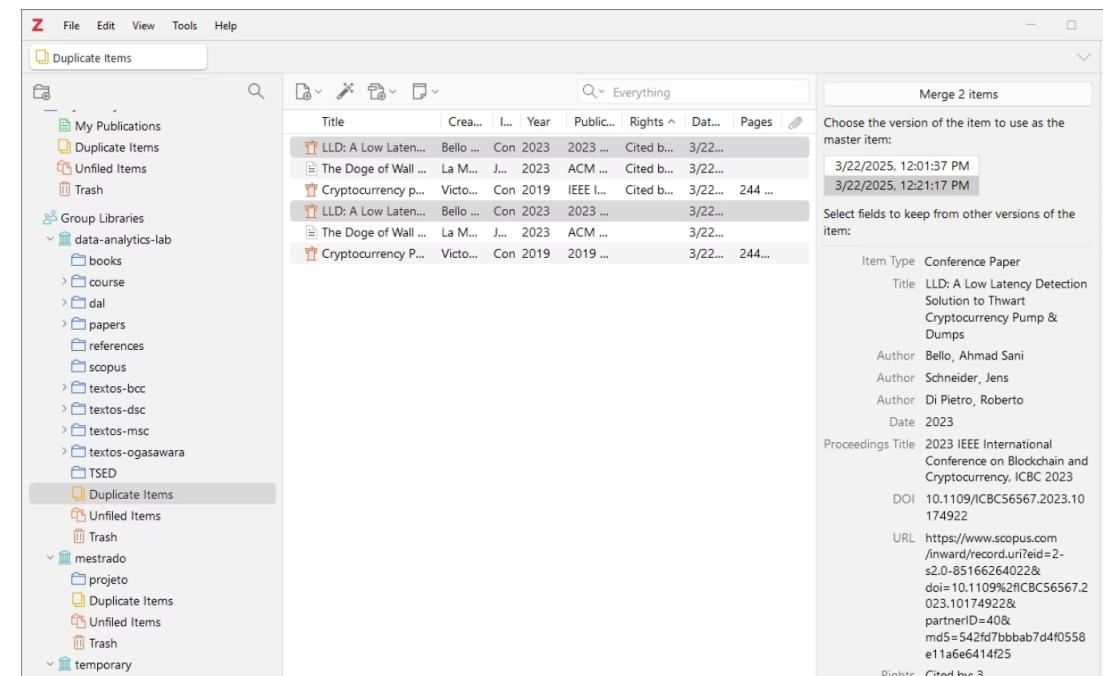
Identificação

Execute novamente a detecção de duplicatas para comparar com referências existentes

Seleção Criteriosa

Escolha preferencialmente os metadados mais recentes e completos vindos do Scopus

Os metadados provenientes do Scopus geralmente são mais atualizados, completos e padronizados, incluindo informações como DOI, número de citações, palavras-chave indexadas e dados completos de publicação.



Durante a mesclagem, preste atenção especial aos campos de autor (formato padronizado), título (capitalização correta), revista (nome oficial completo) e ano de publicação.

Faça o mesmo processo para artigos com título apenas

Nem todas as referências possuem DOI disponível, especialmente artigos mais antigos, capítulos de livros ou publicações em conferências. Para esses casos, utilizamos a busca por título.

"Para os artigos que faltaram curar, busque pelo título com a opção de geração de consulta doi=False"

1 Geração da String

Execute novamente o script R com o parâmetro doi=False para gerar uma string de busca baseada nos títulos das referências ainda não curadas

2 Busca no Scopus

Cole a nova string no campo de busca do Scopus. A busca por título pode retornar mais resultados, então revise cuidadosamente

3 Exportação e Mesclagem

Repita o processo de exportação em formato .bib e importação no Zotero, seguido da identificação e mesclagem de duplicatas

The screenshot shows the Scopus search interface. In the search bar, the query is: TITLE ("A unifying framework for detecting outliers and change points from time series") OR TITLE ("a bitcoin etf is here. what does that mean for investors") OR TITLE ("the way it is and the way it should be: liability under") OR TITLE ("bitcoin trading volumes surge after debut of long-awaited us etfs") OR TITLE ("sequence-based target coin prediction for cryptocurrency pump"). Below the search bar, there are tabs for 'Documents' (selected), 'Preprints', 'Patents', 'Secondary documents', and 'Research data'. The search results section shows '1 document found'. The result is an article by Takeuchi, J.-I. and Yamanishi, K., titled 'A unifying framework for detecting outliers and change points from time series', published in 'IEEE Transactions on Knowledge and Data Engineering', 2006, volume 18(4), pp. 482-492. The document is listed with its title, authors, source, year, and citations.

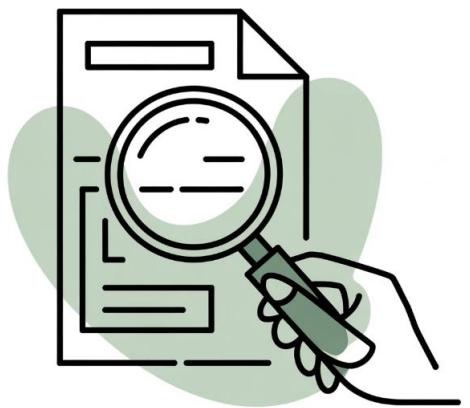
A busca por título exige maior atenção na validação dos resultados, pois pode haver artigos diferentes com títulos similares ou parcialmente correspondentes.

Ajuste manual nas referências que ficaram faltando

Mesmo após realizar buscas por DOI e por título no Scopus, algumas referências podem não ser encontradas. Isso não significa necessariamente que são referências de baixa qualidade, mas merecem atenção especial.

Possíveis Razões

- Artigos publicados em periódicos não indexados no Scopus
- Publicações muito recentes ainda não indexadas
- Livros, relatórios técnicos ou literatura cinzenta
- Documentos em idiomas menos comuns
- Teses e dissertações não disponíveis em bases comerciais



Validação Individual

Verifique cada referência faltante individualmente, buscando em outras bases como IEEE Xplore, ACM Digital Library, PubMed ou Google Scholar

Completude dos Metadados

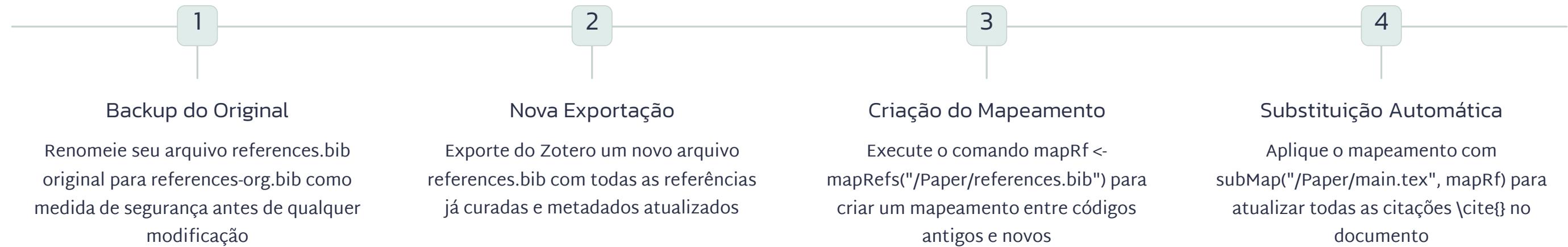
Complete manualmente no Zotero todos os campos disponíveis: autor(es), título completo, local de publicação, editora, volume, número, páginas e DOI quando disponível

Avaliação Crítica

Questione a relevância e qualidade dessas referências. Se não estão em bases reconhecidas, considere substituí-las por fontes mais confiáveis quando possível

Mapeando os códigos antigos para novos

Durante o processo de curadoria, as chaves de citação (citation keys) das referências são modificadas para seguir padrões mais consistentes. É necessário atualizar todas as citações no documento LaTeX.



A screenshot of the RStudio interface. The top menu bar includes File, Edit, Code, View, Plots, Session, Build, Debug, Profile, Tools, Help, and Addins. The main workspace shows an R script named "Untitled1" with the following code:

```
source("https://raw.githubusercontent.com/eogasawara/mylibrary/refs/heads/master/references/myReferences.R")
mapRf <- mapRefs("C:/users/eduar/downloads/Paper/references-org.bib",
                 "C:/users/eduar/downloads/Paper/references.bib")
subMap("C:/users/eduar/downloads/Paper/main.tex", mapRf)
```

The RStudio environment pane shows the Global Environment, History, and Connections tabs. The Files pane shows a folder structure with files like "R816B.Rhistory", "RDData", "Rhitory", "ActivePresenter", "ActivePresenter Templates", "brestfeed", and "Custom Office Templates". The bottom pane shows the Console and Terminal tabs.

- Atenção: Se seu projeto usa múltiplos arquivos .tex (estrutura avançada), execute o comando `subMap()` para cada arquivo que contenha citações bibliográficas.

Limpeza dos metadados desnecessários

O arquivo .bib exportado do Scopus contém muitos campos que não são necessários para a compilação do documento LaTeX e podem causar problemas de formatação. A limpeza remove informações redundantes.

The screenshot shows the RStudio interface. In the top-left corner, there is a script editor window titled "Untitled1" containing R code to clean a BibTeX file. Below it, the "Console" tab displays the output of running this code, showing the number of entries written and the resulting file path. The "Console" tab also shows the R version as "R 4.4.2".

```
source("https://raw.githubusercontent.com/eogasawara/mylibrary/refs/heads/master/references/myReferences.R")
> cleanBib("C:/Users/eduar/Downloads/Paper/references.bib")
writing 23 Bibtex entries ... OK
Results written to file 'C:/users/eduar/Downloads/Paper/references.bib'
>
```

(a) Antes da limpeza

```
#BookShumway_time_2017,
title = {Time Series (Analysis) and (Its) (Applications); (With) (R) (Examples)},
isbn = {978-3-319-54520-9},
shorttitle = {Time Series (Analysis) and (Its) (Applications)},
abstract = {The fourth edition of this popular graduate textbook, like its predecessors, presents a broad treatment of time series analysis for applications in engineering, physical sciences, and business. The authors have focused the text on practical techniques rather than theory, and emphasize model selection, identification, estimation, and forecasting. New to this edition is a chapter on statistical methods with accompanying theory. Numerous examples using nontrivial data illustrate solutions to problems such as discovering natural and anthropogenic climate change, evaluating pain perception experiments using functional magnetic resonance imaging, and monitoring a patient's heart rate. The book also provides R code for each example to allow presentations on different levels. In addition to coverage of classical methods of time series regression, ARIMA models, spectral analysis and state-space models, the text includes modern developments in GARCH models, Bayesian models, state-space models, stochastic volatility, wavelets, and Harary chain Monte Carlo integration methods. This edition includes R code for each example which can be run directly from the book's website for download, making all the data sets and scripts easy to load into R.},
language = {English},
note = {Scholar: 78751},
pubdate = {2017-04-01},
author = {Shumway, Robert H. and Stoffer, David S.},
month = {april},
year = {2017},
doi = {10.1007/978-3-319-54520-9}
```

Arquivo com campos excessivos como abstract, keywords, note, url, issn, que ocupam espaço e raramente são usados na bibliografia final

(b) Depois da limpeza

```
1 #Book(shumway_time_2017,
2 title = {Time Series (Analysis) and (Its) (Applications); (With) (R) (Examples)},
3 isbn = {978-3-319-54520-9},
4 shorttitle = {Time Series (Analysis) and (Its) (Applications)},
5 language = {en},
6 publisher = {Springer},
7 author = {Robert H. Shumway and David S. Stoffer},
8 month = {april},
9 year = {2017},
10 )
11
12 #BookKhan_data_2022,
13 title = {Data (Mining): (Concepts) and (Techniques)},
14 isbn = {978-3-030-74760-6},
15 shorttitle = {Data (Mining)},
16 language = {English},
17 publisher = {Morgan Kaufmann},
18 author = {Jiawei Han and Jian Pei and Hanghang Tong},
19 month = {oct},
20 year = {2022},
21 address = {Cambridge, MA},
22 edition = {4th edition},
23 }
```

Arquivo otimizado mantendo apenas os campos essenciais: author, title, journal, year, volume, pages e doi

Comando de Limpeza

```
cleanBib("/Paper/references.bib")
```

Execute este comando simples para remover automaticamente campos desnecessários

Benefícios

- Reduz significativamente o tamanho do arquivo
- Facilita leitura e manutenção manual quando necessário
- Evita conflitos com estilos bibliográficos específicos
- Melhora performance de compilação em documentos grandes

Integração com Overleaf



Após concluir todo o processo de curadoria localmente, é hora de integrar os arquivos atualizados com o ambiente de edição colaborativa Overleaf.

Processo de Integração

1. Acesse seu projeto no Overleaf através do navegador
2. Faça upload dos arquivos .tex atualizados (com citações corrigidas)
3. Substitua o arquivo references.bib antigo pelo novo arquivo curado e limpo
4. Compile o documento para verificar se todas as referências foram corretamente atualizadas
5. Verifique a seção de referências bibliográficas no PDF gerado

1

Verificação Final

Certifique-se de que todas as citações no texto correspondem às entradas na bibliografia e que não há avisos de referências indefinidas

2

Colaboração

Compartilhe o projeto atualizado com seus orientadores e coautores, que agora terão acesso a referências padronizadas e confiáveis

3

Manutenção Contínua

Sempre que adicionar novas referências, repita o processo de curadoria para manter a qualidade e consistência da bibliografia

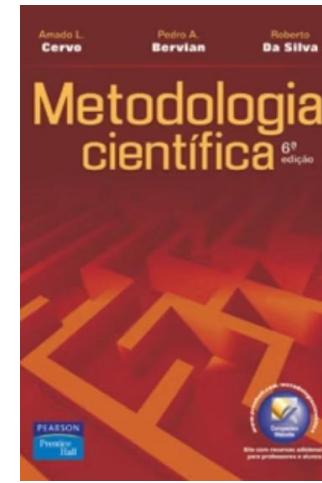
Parabéns! Você concluiu o processo completo de curadoria de referências. Sua pesquisa agora possui uma base bibliográfica sólida, confiável e reproduzível, elevando significativamente a qualidade acadêmica do seu trabalho.

Referências Bibliográficas

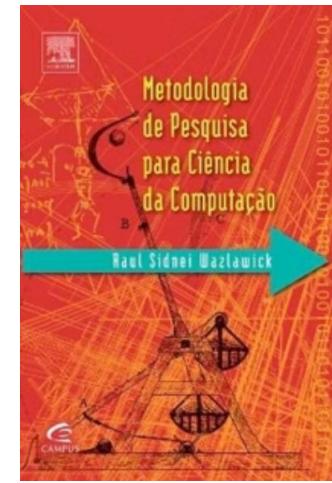
Esta apresentação foi desenvolvida com base em obras fundamentais sobre metodologia científica e escrita acadêmica, essenciais para o desenvolvimento de competências em pesquisa e análise de artigos científicos.



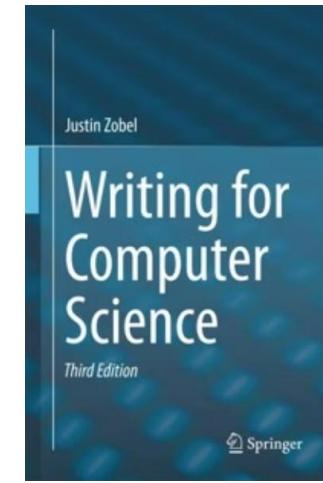
Perovano (2016)
Manual de metodologia da pesquisa científica - Editora Intersaberes.
Obra completa sobre fundamentos metodológicos.



Cervo, Bervian & Silva (2006)
Metodologia Científica - Pearson Universidades. Referência clássica em metodologia de pesquisa.



Wazlawick (2017)
Metodologia de Pesquisa para Ciência da Computação - Elsevier Brasil. Específico para área de computação.



Zobel (2015)
Writing for Computer Science - Springer. Guia essencial para escrita científica em computação.