

# Citações de Trabalhos e Plágio



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

#### O que são citações?

- Citação é a menção no texto de informações extraídas de outras fontes
- Elas situam o trabalho no contexto da temática, conferindo credibilidade
- Permitem a confrontação de dados, fatos e argumentos
- Devem ser sólidas e bem localizadas no texto
- Evite citar apenas para aumentar a bibliografia
- Exemplo de citação correta:
  - A aprendizagem contínua é essencial para a adaptação de modelos de IA (Zobel, 2015).
- Exemplo de citação ruim:
  - Pesquisadores já discutiram a importância da aprendizagem contínua (Zobel, 2015, Wazlawick, 2017, Smith, 2014, Jones, 2016, ...).
    - Muitas referências irrelevantes sem contextualização

### Critérios para incluir uma citação

#### Cite:

- Artigos originais ao invés de fontes secundárias
- Trabalhos bem escritos e revisados
- Documentos formais como white papers
- X Não cite:
  - Palestras ou seminários (se essencial, inclua como nota de rodapé)
  - Manuscritos não revisados
  - Informações de senso comum (ex.: Wikipedia sem fontes verificáveis).
- Exemplo de boa prática:
  - "Segundo Zobel (2015), a escrita científica deve ser direta e objetiva."
    - Citação de um livro com credibilidade.
- Exemplo de má prática:
  - "Segundo um artigo que li na internet, a escrita científica deve ser objetiva."
    - Falta de fonte confiável.

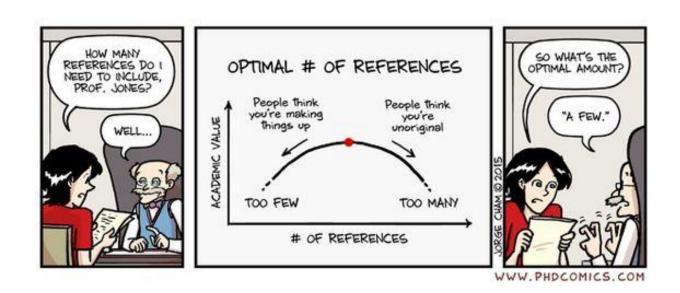
#### Autocitação: Quando Usar e Quando Evitar?

#### Use quando:

- Seus trabalhos anteriores são relevantes para o novo estudo
- Demonstram continuidade da pesquisa
- Evite quando:
  - O objetivo é apenas inflar as métricas de citação
  - O trabalho anterior não tem relação direta com o novo
- Exemplo positivo:
  - No estudo anterior (Silva et al., 2020), testamos um modelo estatístico. Neste trabalho, expandimos o modelo considerando redes neurais.
- Exemplo negativo:
  - Nossa pesquisa (Silva et al., 2020) já tratou desse tema. Para mais detalhes, veja Silva et al. (2018), Silva et al. (2017), Silva et al. (2015).
    - Autocitação excessiva sem justificativa

### Onde as citações são mais frequentes?

- Introdução → Para situar o problema
- Referencial teórico → Para embasar conceitos
- Trabalhos relacionados → Para comparar abordagens



#### Formatos de citações e referências: Autor-Ano

- 1. Autor-Ano (ABNT, SBC)
- Ex.: Segundo Wazlawick (2017), a pesquisa científica deve seguir padrões éticos.



#### Referências

- J. Gama, I. Zliobaite, A. Bifet, M. Pechenizkiy, and A. Bouchachia. A survey on concept drift adaptation. ACM Computing Surveys, 46(4), 2014.
- Damodar N. Gujarati and Dawn C. Porter. Basic Econometrics. McGraw-Hill Publishing, 2008. ISBN 978-0-07-741480-1.
- T.R. Hoens, R. Polikar, and N.V. Chawla. Learning from streaming data with concept drift and imbalance: An overview. Progress in Artificial Intelligence, 1(1):89–101, 2012.
- A.S. Iwashita and J.P. Papa. An Overview on Concept Drift Learning. IEEE Access, 7:1532–1547, 2019.
- I. Khamassi, M. Sayed-Mouchaweh, M. Hammami, and K. Ghédira. Discussion and review on evolving data streams and concept drift adapting. Evolving Systems, 9(1), 2018.
- J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang. Learning under Concept Drift: A Review. IEEE Transactions on Knowledge and Data Engineering, 31(12):2346–2363, 2019.
- S. Wang, L.L. Minku, and X. Yao. A Systematic Study of Online Class Imbalance Learning with Concept Drift. IEEE Transactions on Neural Networks and Learning Systems, 29(10):4802–4821, 2018.
- G.I. Webb, R. Hyde, H. Cao, H.L. Nguyen, and F. Petitjean. Characterizing concept drift. Data Mining and Knowledge Discovery, 30(4):964–994, 2016.

Em contextos de conjunto de dados multidimensionais, a não estacionariedade fica bem caracterizada a partir do desvio de conceito (do inglês, concept drift) [Iwashita and Papa, 2019]. Considere o problema geral de predição, de modo que se tem um conjunto de variáveis endógenas (explicativas) X usadas para predizer a variável exógena (de resposta) Y [Gujarati and Porter, 2008]. O desvio de conceito é caracterizado como sendo uma mudança significativa na distribuição de dados ou no relacionamento entre a variável de saída Y e as variáveis de entrada X [Webb et al., 2016; Wang et al., 2018]. Formalmente, um conceito no tempo i é definido como sendo a probabilidade conjunta  $\chi$  de X e Y, descrita pela Equação 1. Um desvio de conceito entre os tempos i e j é definido como sendo uma diferença significativa entre os conceitos  $p_4(\chi)$  e  $p_3(\chi)$ , como descrito pela Equação 2 [Gama et al., 2014; Khamassi et al., 2018].

$$concept_i = p_i(\chi) = p_i(X, Y)$$
 (1)

$$p_i(\chi) \neq p_j(\chi)$$
 (2)

O desvio de conceito pode ser classificado como real ou virtual [Iwashita and Papa, 2019]. Especificamente, o desvio de conceito real é definido como sendo uma mudança nas bordas que caracterizam a variável de saída (probabilidade a posteriori), i.e., p(Y|X). Reciprocamente, no desvio de conceito virtual tem-se que as bordas que caracterizam a variável de saída são mantidas. Neste caso, há uma mudança na probabilidade da variável de saída (p(Y)) ou na probabilidade condicional (p(X|Y)) ao longo do tempo [Iwashita and Papa, 2019; Lu et al., 2019; Hoens et al., 2012].

A Figura 1.a apresenta um classificador treinado para um conjunto de dados com duas classes:  $Y_1$  e  $Y_2$ . O classificador estabelece uma superfície de separação para os dados. Na Figura 1.b, tem-se desvio de conceito virtual, onde a superfície de separação entre as classes não muda, mas a relação de frequência entre as classes  $Y_1$  e  $Y_2$  é diferente. Há mais observações associadas ao rótulo  $Y_2$  do que ao rótulo  $Y_1$  quando comparado ao cenário inicial. Já na Figura 1.c, tem-se o desvio de conceito real, onde a superfície de alteração anterior já não é adequada. Nota-se que as posições no espaço de atributos para as observações associadas ao rótulo  $Y_1$  se aproximam daquelas observações associadas ao rótulo  $Y_2$ , embora se tenha a mesma quantidade de observações em relação ao cenário inicial.

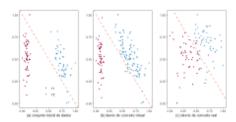


Figura 1. (a) Classificador para  $Y_1$  e  $Y_2$ ; (b) desvio virtual; (c) desvio real

#### Formatos de citações e referências: Numérico

- Numérico (IEEE, ACM)
- Ex.: A pesquisa científica deve seguir padrões éticos [1].



#### Referências

- J. Gama, I. Zliobaite, A. Bifet, M. Pechenizkiy, and A. Bouchachia. A survey on concept drift adaptation. ACM Computing Surveys, 46(4), 2014.
- [2] Damodar N. Gujarati and Dawn C. Porter. Basic Econometrics. McGraw-Hill Publishing, 2008. ISBN 978-0-07-741480-1.
- [3] T.R. Hoens, R. Polikar, and N.V. Chawla. Learning from streaming data with concept drift and imbalance: An overview. Progress in Artificial Intelligence, 1(1):89–101, 2012.
- [4] A.S. Iwashita and J.P. Papa. An Overview on Concept Drift Learning. IEEE Access, 7:1532–1547, 2019.
- [5] I. Khamassi, M. Sayed-Mouchaweh, M. Hammami, and K. Ghédira. Discussion and review on evolving data streams and concept drift adapting. Evolving Systems, 9(1), 2018.
- [6] J. Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang. Learning under Concept Drift: A Review. IEEE Transactions on Knowledge and Data Engineering, 31(12):2346–2363, 2019.
- [7] S. Wang, L.L. Minku, and X. Yao. A Systematic Study of Online Class Imbalance Learning with Concept Drift. IEEE Transactions on Neural Networks and Learning Systems, 29(10):4802–4821, 2018.
- [8] G.I. Webb, R. Hyde, H. Cao, H.L. Nguyen, and F. Petitjean. Characterizing concept drift. Data Mining and Knowledge Discovery, 30(4):964–994, 2016.

Em contextos de conjunto de dados multidimensionais, a não estacionariedade fica bem caracterizada a partir do desvio de conceito (do inglês, concept drift) [4]. Considere o problema geral de predição, de modo que se tem um conjunto de variáveis endógenas (explicativas) X usadas para predizer a variável exógena (de resposta) Y [2]. O desvio de conceito é caracterizado como sendo uma mudança significativa na distribuição de dados ou no relacionamento entre a variável de saída Y e as variáveis de entrada X [8, 7]. Formalmente, um conceito no tempo i é definido como sendo a probabilidade conjunta  $\chi$  de X e Y, descrita pela Equação 1. Um desvio de conceito entre os tempos i e j é definido como sendo uma diferença significativa entre os conceitos  $p_i(\chi)$  e  $p_j(\chi)$ , como descrito pela Equação 2 [1, 5].

$$concept_t = p_t(\chi) = p_t(X, Y)$$
 (1)

$$p_i(\chi) \neq p_j(\chi)$$
 (2)

O desvio de conceito pode ser classificado como real ou virtual [4]. Especificamente, o desvio de conceito real é definido como sendo uma mudança nas bordas que caracterizam a variável de saída (probabilidade a posteriori), i.e., p(Y|X). Reciprocamente, no desvio de conceito virtual tem-se que as bordas que caracterizam a variável de saída são mantidas. Neste caso, há uma mudança na probabilidade da variável de saída (p(Y)) ou na probabilidade condicional (p(X|Y)) ao longo do tempo [4, 6, 3].

A Figura 1.a apresenta um classificador treinado para um conjunto de dados com duas classes:  $Y_1$  e  $Y_2$ . O classificador estabelece uma superfície de separação para os dados. Na Figura 1.b, tem-se desvio de conceito virtual, onde a superfície de separação entre as classes não muda, mas a relação de frequência entre as classes  $Y_1$  e  $Y_2$  é diferente. Há mais observações associadas ao rótulo  $Y_2$  do que ao rótulo  $Y_1$  quando comparado ao cenário inicial. Já na Figura 1.c, tem-se o desvio de conceito real, onde a superfície de alteração anterior já não é adequada. Nota-se que as posições no espaço de atributos para as observações associadas ao rótulo  $Y_2$ , embora se tenha a mesma quantidade de observações em relação ao cenário inicial.

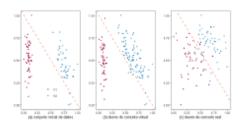


Figura 1. (a) Classificador para Y1 e Y2; (b) desvio virtual; (c) desvio real

#### Dimensão de ideia-autoridade

- Citação por ideia:
  - A referência aparece entre colchetes no final
  - Exemplo: O desvio de conceito pode ser classificado como real ou virtual [4].
- Citação por autoridade:
  - O nome do autor faz parte do texto
  - Exemplo: Iwashita e Papa [2019] classificam o desvio de conceito como real ou virtual.

#### Citações diretas e indiretas

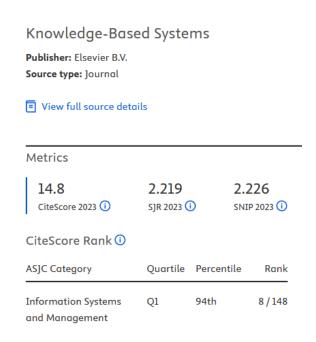
- Citação Direta (uso de aspas)
  - "A escrita científica deve ser clara e objetiva" (Zobel, 2015).
- Citação Indireta (reformulação com suas palavras)
  - Zobel (2015) afirma que a escrita científica deve ser clara e objetiva.
- Evite paráfrases sem citação. Isso pode ser considerado plágio

#### Como pesquisar referências?

- Use fontes científicas indexadas:
  - Scopus, IEEE Xplore, ACM Digital Library
- Fóruns relevantes (JCR, SJR)
- Artigos com impacto (Citações, FWCI)
- Evite citar blogs, Wikipedia e fontes sem credibilidade acadêmica

## Nonstationary time series transformation methods: An experimental review





#### Coerência entre a seção de referências e texto

- Tudo o que é citado deve aparecer listado nas referências bibliográficas
- Tudo o que aparece nas referências bibliográficas tem que ter sido citado no texto
- Não é produtivo controlar as referências bibliográficas "na mão"
- Use sistemas de gerência de referências

Em contextos de conjunto de dados multidimensionais, a não estacionariedade fica bem caracterizada a partir do desvio de conceito (do inglês, concept drifi) [Iwashita and Papa, 2019]. Considere o problema geral de predição, de modo que se tem um conjunto de variáveis endógenas (explicativas) X usadas para predizer a variável exógena (de resposta) Y [Gujarati and Porter, 2008]. O desvio de conceito é caracterizado como sendo uma mudança significativa na distribuição de dados ou no relacionamento entre a variável de saída Y e as variáveis de entrada X [Webb et al., 2016; Wang et al., 2018]. Formalmente, um conceito no tempo i é definido como sendo a probabilidade conjunta  $\chi$  de X e Y, descrita pela Equação 1. Um desvio de conceito entre os tempos i e j é definido como sendo uma diferença significativa entre os conceitos  $p_i(\chi)$  e  $p_j(\chi)$ , como descrito pela Equação 2 [Gama et al., 2014; Khamassi et al., 2018].

$$concept_i = p_i(\chi) = p_i(X, Y)$$
 (1)  
 $p_i(\chi) \neq p_i(\chi)$  (2)

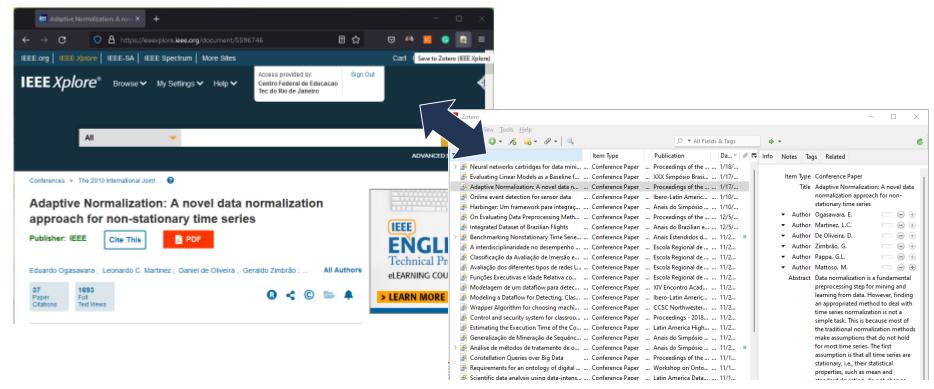
O desvio de conceito pode ser classificado como real ou virtual [Iwashita and Papa, 2019]. Especificamente, o desvio de conceito real é definido como sendo uma mudança nas bordas que caracterizam a variável de saída (probabilidade a posteriori), i.e., p(Y|X). Reciprocamente, no desvio de conceito virtual tem-se que as bordas que caracterizam a variável de saída são mantidas. Neste caso, há uma mudança na probabilidade da variável de saída (p(Y)) ou na probabilidade condicional (p(X|Y)) ao longo do tempo [Iwashita and Papa, 2019; Lu et al., 2019; Hoens et al., 2012].

#### Referências

- J. Gama, I. Zliobaite, A. Bifet, M. Pechenizkiy, and A. Bouchachia. A survey on concept drift adaptation. ACM Computing Surveys, 46(4), 2014.
- Damodar N. Gujarati and Dawn C. Porter. Basic Econometrics. McGraw-Hill Publishing, 2008. ISBN 978-0-07-741480-1.
- T.R. Hoens, R. Polikar, and N.V. Chawla. Learning from streaming data with concept drift and imbalance: An overview. Progress in Artificial Intelligence, 1(1):89–101, 2012.
- A.S. Iwashita and J.P. Papa. An Overview on Concept Drift Learning. IEEE Access, 7:1532–1547, 2019.
- Khamassi, M. Sayed-Mouchaweh, M. Hammami, and K. Ghédira. Discussion and review on evolving data streams and concept drift adapting. Evolving Systems, 9(1), 2018.
- Lu, A. Liu, F. Dong, F. Gu, J. Gama, and G. Zhang. Learning under Concept Drift: A Review. IEEE Transactions on Knowledge and Data Engineering, 31(12):2346–2363, 2019.
- S. Wang, L.L. Minku, and X. Yao. A Systematic Study of Online Class Imbalance Learning with Concept Drift. IEEE Transactions on Neural Networks and Learning Systems, 29(10):4802–4821, 2018.
- G.I. Webb, R. Hyde, H. Cao, H.L. Nguyen, and F. Petitjean. Characterizing concept drift. Data Mining and Knowledge Discovery, 30(4):964–994, 2016.

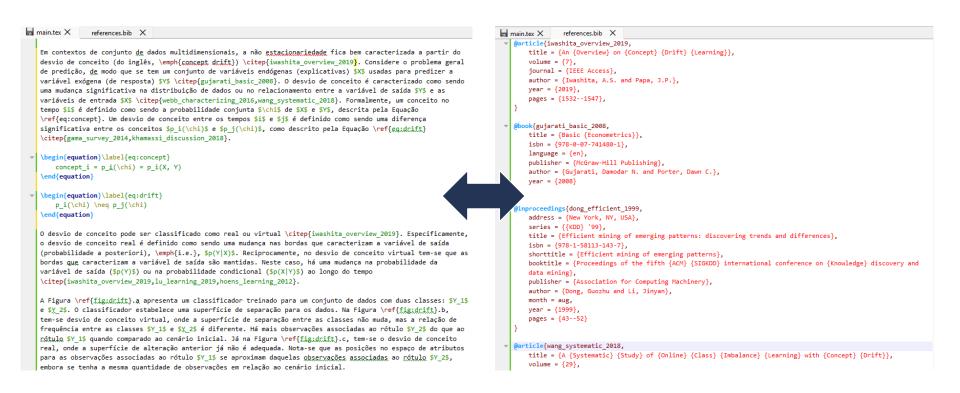
#### Gerenciamento de referências

- Use ferramentas para evitar erros manuais:
  - Zotero, Mendeley, JabRef
- Integre com editores como LaTeX e Word



#### Editores preparados para trabalhar com referências

- Editores de textos científicos, como o LaTeX, já são preparados para trabalhar bem com referências
- Basta incluir um arquivo BibTeX no projeto do documento e colocar referências cruzadas no texto (veja slide de Introdução ao LaTeX)



## Plágio

### Plágio: Definição e Consequências

- Plágio ocorre quando um autor usa o trabalho de outra pessoa sem dar os devidos créditos
- Tipos de Plágio:
  - Acidental Falta de conhecimento sobre normas
  - Não Intencional Ideias semelhantes surgem naturalmente
  - Intencional Cópia proposital de um trabalho
  - Autoplágio Uso de trabalho próprio sem citação

## Ferramentas de Detecção de Plágio

- Ferramentas
  - CopySpider, Plagium, Copyscape
  - Plagius (português)
- Revisores são treinados para identificar plágio

### Código de ética e regras sobre plágio

- ACM Code of Ethics:
  - "One must not take credit for other's ideas or work."
- IEEE Code of Ethics:
  - "To credit properly the contributions of others."
- Código de Conduta SBC:
  - Plágio é considerado fraude
  - Autoplágio é antiético
  - Submissão múltipla sem permissão é inaceitável

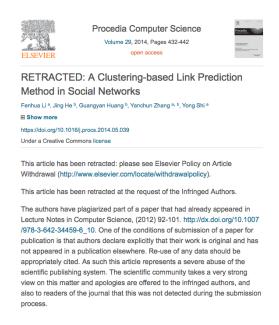
#### Retratação de artigos por plágio

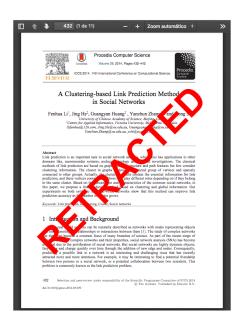
- A Consequências graves para o pesquisador
  - Perda de credibilidade acadêmica
  - Dificuldade em publicar novos artigos
  - Penalizações por instituições acadêmicas e agências de fomento



#### Retratação de artigo por autoplágio

- Recomendação:
  - Se precisar reutilizar conteúdo de artigos anteriores, sempre cite explicitamente a fonte e deixe claro quais partes foram reaproveitadas
  - Evite submeter o mesmo artigo a múltiplos periódicos sem autorização explícita

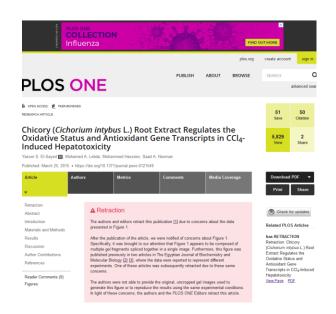




### Retratação de artigos por plágio em figuras

- Plágio ocorre quando imagens são reutilizados sem permissão ou citação
- Recomendação:
  - Sempre cite a fonte de figuras extraídas de outros trabalhos
  - Use imagens de domínio público ou peça permissão ao autor original
  - Evite manipular imagens para alterar a interpretação dos dados





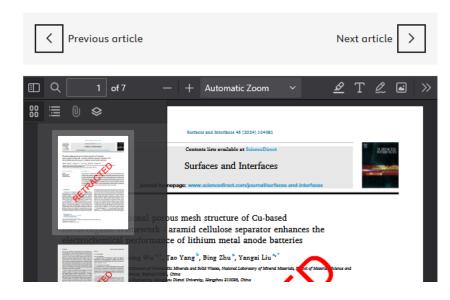
#### Retratação por conta de uso de Large Language Models (LLM)

- Problemas causados pelo uso de IA na escrita científica:
  - Erros factuais: Modelos de IA podem gerar informações imprecisas
  - Conteúdo gerado sem revisão: Possibilidade de referências falsas
  - Falta de transparência: Não declarar o uso de IA pode ser considerado antiético

An investigation by the journal confirmed substantial duplication of text and image data between these two articles that were submitted and published in close succession. All authors of the International Journal of Hydrogen Energy article also authored the Surfaces and Interfaces article.

In addition, there are concerns that the authors appear to have used a Generative AI source in the writing process of the paper without disclosure, which is a breach of journal policy.

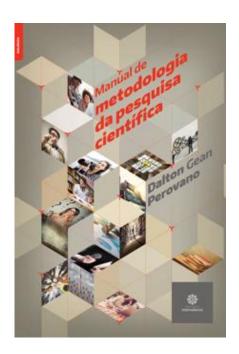
The journal sincerely regrets that these issues were not detected during the manuscript screening and evaluation process and apologies are offered to readers of the journal.

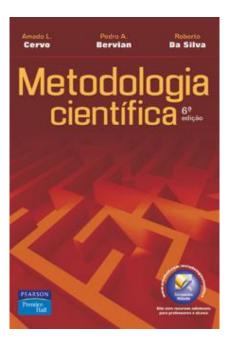


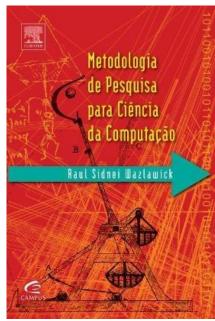
#### Resumo Final

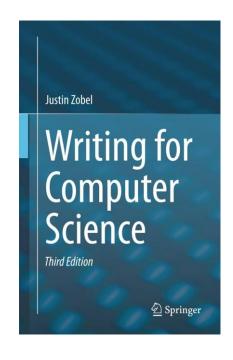
- Argumente com suas próprias palavras
- Cite corretamente todas as referências
- Nunca copie trechos sem citação
- Use ferramentas de referência e detecção de plágio
- Boa escrita científica começa com boas práticas de citação

### Referências









[1] D. G. Perovano, Manual de metodologia da pesquisa científica. Editora Intersaberes, 2016. [2] A. L. Cervo, P. A. Bervian, e R. da Silva, Metodologia Científica. Pearson Universidades, 2006. [3] R. Wazlawick, 2017, Metodologia de Pesquisa para Ciência da Computação. Elsevier Brasil. [4] J. Zobel, 2015, Writing for Computer Science. Springer.

