



Pesquisa em Computação

Uma jornada pelos métodos, abordagens e estratégias que fundamentam a produção científica e tecnológica na área de Computação.

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

Duas Faces da Mesma Moeda

Ciência

A ciência busca conhecimento e explicações sobre o mundo. Seu principal objetivo é construir teorias robustas que expliquem fatos observados e prevejam fenômenos futuros. Na computação, trabalhos científicos se dedicam a explicar **como e por que** algo funciona, explorando princípios fundamentais.



Exemplo Científico

A **conjectura do favo de mel** demonstra matematicamente por que estruturas hexagonais são ótimas para dividir superfícies com o mínimo perímetro possível.

Tecnologia

A tecnologia aplica conhecimento para resolver problemas práticos do cotidiano. Seu foco não é explicar fenômenos naturais, mas desenvolver soluções que impactam o mundo real. Trabalhos tecnológicos resultam em ferramentas, algoritmos e processos que transformam teoria em prática.



Exemplo Tecnológico

O **H3 do Uber** aplica essa teoria em uma estrutura espacial hexagonal hierárquica para indexação geográfica eficiente em aplicações reais.

- **Referências:**[1] T.C. Hales, 2001, The Honeycomb Conjecture, *Discrete & Computational Geometry*, v. 25, n. 1, p. 1–22.[2] I. Brodsky, 2018. H3: Uber's Hexagonal Hierarchical Spatial Index. *Uber Engineering Blog*. <https://eng.uber.com/h3>

Classificando Abordagens de Investigação

A pesquisa em Computação pode ser classificada de acordo com diferentes dimensões que caracterizam sua natureza e metodologia. Compreender essas classificações é fundamental para escolher a abordagem adequada ao problema investigado.

Fonte de Dados

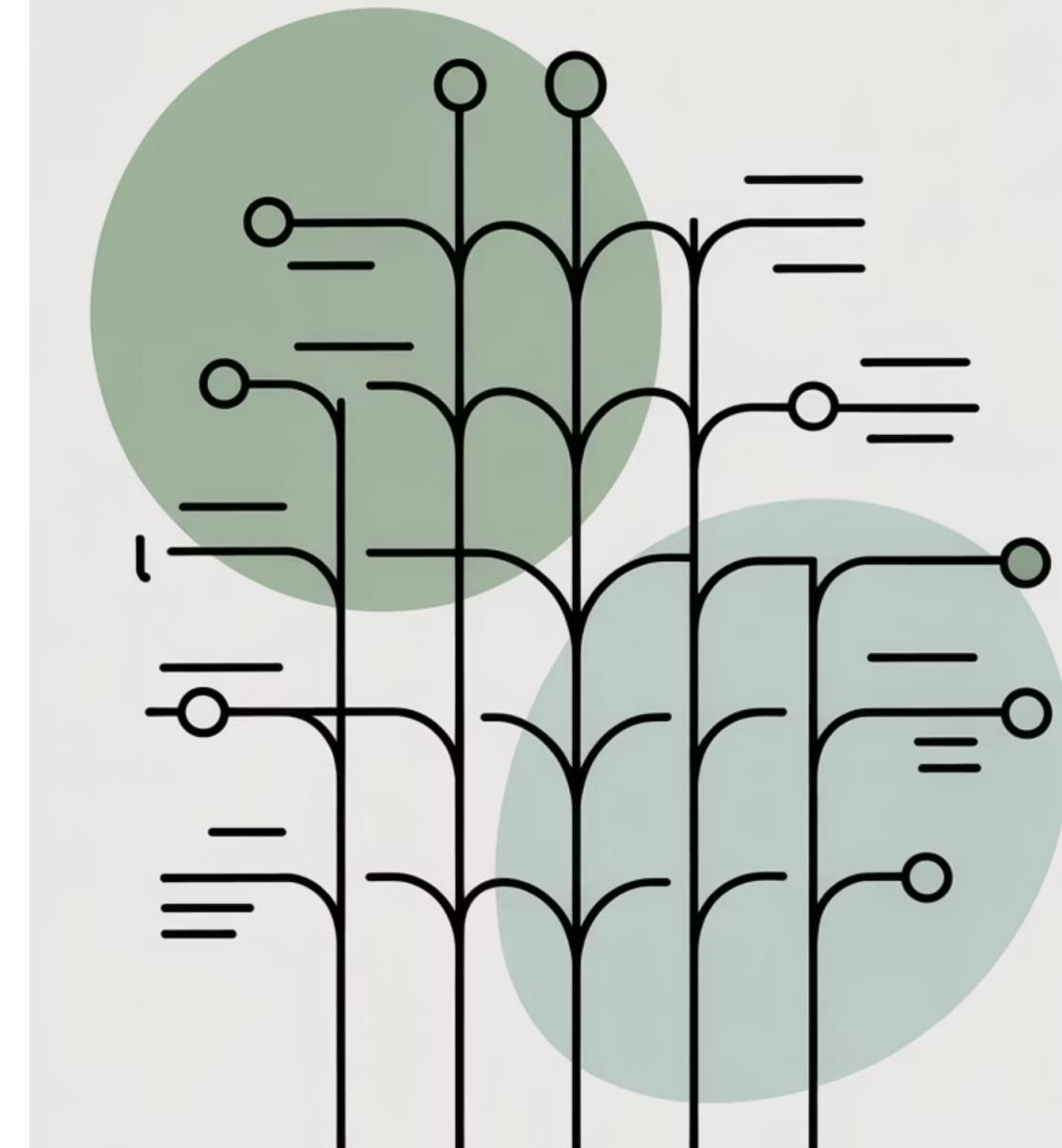
Define a origem das informações utilizadas na pesquisa:

- **Primária:** Produção de novos dados através de experimentos, entrevistas ou observações diretas
- **Secundária:** Sistematização e análise de conhecimento já existente na literatura

Métodos de Investigação

Caracteriza como os dados são coletados e analisados:

- **Quantitativa:** Utiliza dados numéricos e métodos estatísticos
- **Qualitativa:** Trabalha com dados textuais e análises descritivas
- **Analítica:** Constrói teorias baseadas em provas matemáticas formais



Pesquisa Primária vs. Secundária

Produzindo ou Sistematizando Conhecimento

Pesquisa Primária

Produz **novos dados** através de experimentos controlados, entrevistas estruturadas ou observações sistemáticas. Esta abordagem é fundamental para expandir as fronteiras do conhecimento.

Metodologias aplicáveis:

- Qualitativa
- Quantitativa
- Analítica

Exemplos práticos:

- Desenvolvimento de um novo algoritmo de aprendizado de máquina
- Criação de um novo método de normalização para séries temporais

Pesquisa Secundária

Sistematiza o **conhecimento existente** através da análise crítica e síntese da literatura. Identifica lacunas, tendências e oportunidades de pesquisa futura.

Principais formatos:

- Revisões sistemáticas da literatura
- Meta-análises quantitativas
- Mapeamentos sistemáticos

Exemplos práticos:

- Revisão sistemática sobre técnicas de aprendizado de máquina aplicadas a problemas específicos
- Meta-análise de métodos de normalização em diferentes domínios



Pesquisa Quantitativa vs. Qualitativa

Números ou Narrativas?

Pesquisa Quantitativa

Trabalha com **dados numéricos** e utiliza métodos estatísticos rigorosos para testar hipóteses e validar modelos. A objetividade e reproduzibilidade são características essenciais.

Características principais:

- Coleta de dados mensuráveis
- Análise estatística e inferencial
- Testes de hipóteses formais
- Resultados generalizáveis

Exemplo concreto:

Avaliação da acurácia de um modelo de IA utilizando métricas quantitativas como precisão, recall, F1-score e AUC-ROC em datasets padronizados.

Pesquisa Qualitativa

Trabalha com **dados textuais ou subjetivos**, utilizando métodos descritivos e exploratórios para compreender fenômenos complexos. O foco está na profundidade da compreensão.

Características principais:

- Coleta de dados narrativos
- Análise temática e interpretativa
- Compreensão de contextos
- Insights profundos sobre comportamentos

Exemplo concreto:

Entrevistas semiestruturadas com **desenvolvedores de software** para compreender os desafios, motivações e barreiras na adoção de uma nova tecnologia ou framework em suas equipes.

Pesquisa Analítica

O Rigor da Matemática

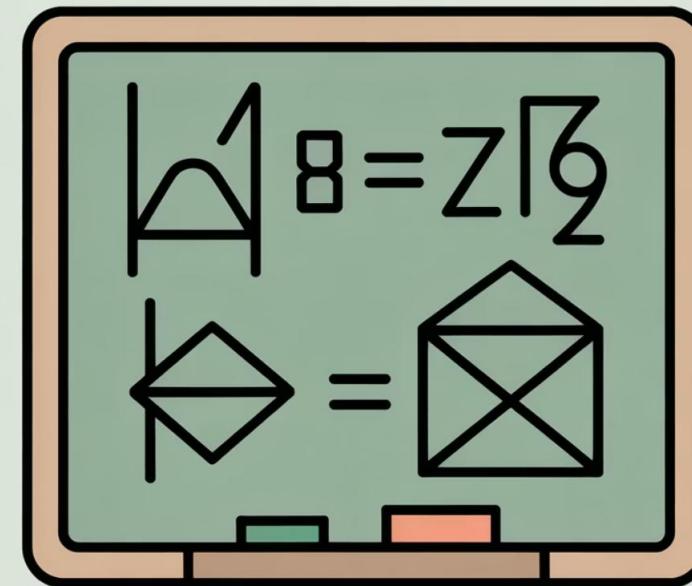
A pesquisa analítica constrói **teorias formais** e apresenta **provas matemáticas rigorosas** que garantem a correção e as propriedades de algoritmos e sistemas computacionais.

Esta abordagem é particularmente comum em áreas fundamentais da computação que exigem garantias formais de comportamento e desempenho.

Áreas de aplicação típicas:

- **Criptografia:** Provas de segurança e complexidade
- **Teoria da Computação:** Limites de computabilidade
- **Análise de Algoritmos:** Complexidade temporal e espacial
- **Lógica Formal:** Verificação de programas

- ❑ **Exemplo clássico:** Demonstração formal de que o algoritmo de Dijkstra para encontrar o caminho mais curto tem complexidade $O(V^2)$ com lista de adjacência simples, ou $O((V + E) \log V)$ com heap binário.



Comparação e Validação de Contribuições

Ao apresentar uma nova contribuição científica em Computação, é fundamental contextualizá-la adequadamente e avaliar sua validade de forma rigorosa. A estratégia de validação depende diretamente da natureza do trabalho realizado.

01

Algo Diferente

Quando não há trabalhos semelhantes, construa uma argumentação sólida sobre o valor e impacto da contribuição

02

Algo Melhor

Compare sua solução com benchmarks reconhecidos para demonstrar superioridade

03

Uma Prova

Em áreas teóricas, valide através de demonstrações matemáticas formais

04

Um Artefato

Quando o foco é software, avalie através de testes práticos e estudos de caso

05

Artigos de Dados

Descreva claramente o impacto e potencial de reuso do dataset organizado

Pergunta-chave fundamental:

"Como posso demonstrar convincentemente o valor e a originalidade da minha pesquisa para a comunidade científica?"

Tipo 1: Apresentação de Algo Diferente

Pioneirismo sem Precedentes

Este tipo de pesquisa aborda problemas ou soluções **sem trabalhos diretamente relacionados** na literatura. A ausência de comparações diretas torna a validação mais desafiadora, exigindo criatividade metodológica.

Características principais:

- Avaliação predominantemente qualitativa
- Uso extensivo de **estudos de caso** para evidências empíricas
- Argumentação forte sobre relevância e impacto
- Análise de viabilidade e aplicabilidade

Exemplo ilustrativo:

Desenvolvimento de um **novo** paradigma de programação que não se enquadra nas categorias existentes (imperativo, funcional, lógico, orientado a objetos), exigindo validação através de casos de uso práticos.



Contexto

Estabeleça o problema e justifique por que abordagens existentes são insuficientes

Solução Proposta

Descreva detalhadamente sua contribuição original e seus diferenciais

Argumentação

Demonstre valor através de estudos de caso, análises qualitativas e discussão de benefícios

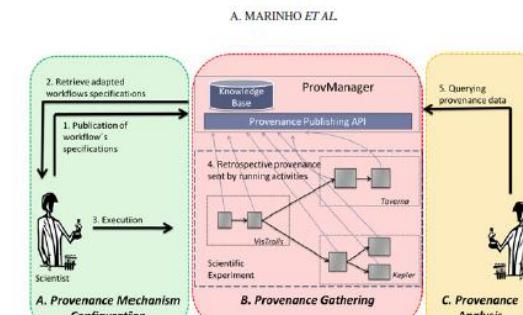


Figure 1. ProvManager in operation.

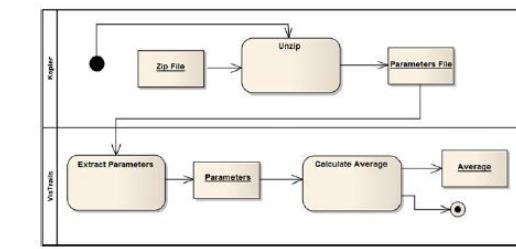


Figure 11. Diagram of CHExp.

[1] A. Marinho, L. Murta, C. Werner, V. Braganholo, S.M.S.D. Cruz, E. Ogasawara, and M. Mattoso, 2012, ProvManager: A provenance management system for scientific workflows, *Concurrency and Computation: Practice and Experience*, v. 24, n. 13, p. 1513–1530.

Superando o Estado da Arte

Quando existem soluções prévias para o problema investigado, a nova abordagem deve ser **rigorosamente comparada** com trabalhos relacionados para demonstrar suas vantagens. Esta é a forma mais comum de validação em Computação.

Estratégias de comparação:

- **Uso de benchmarks estabelecidos:** Datasets e problemas padrão reconhecidos pela comunidade (ex: MNIST, ImageNet, COCO)
- **Criação de experimentos controlados:** Quando benchmarks não existem, o autor pode criar testes específicos, mas deve ter cuidado para evitar viés
- **Definição de métricas claras:** Estabelecer critérios objetivos de avaliação (acurácia, tempo de execução, uso de memória)
- **Testes estatísticos:** Validar significância das diferenças observadas



Identifique Baselines

Selecione trabalhos representativos do estado da arte como pontos de comparação



Execute Experimentos

Teste sua solução e as baselines nas mesmas condições controladas



Analise Resultados

Compare métricas quantitativas e discuta vantagens e limitações

- Exemplo:** Um novo algoritmo de aprendizado de máquina testado em datasets padrão como MNIST e ImageNet, comparando acurácia, tempo de treinamento e robustez com técnicas como ResNet, VGG e EfficientNet.

[1] R. Salles, P. Mattos, A.-M.D. Iorgulescu, E. Bezerra, L. Lima, and E. Ogasawara, 2016, Evaluating temporal aggregation for predicting the sea surface temperature of the Atlantic Ocean, *Ecological Informatics*, v. 36, p. 94–105.

Tipo 3: Apresentação de Uma Prova

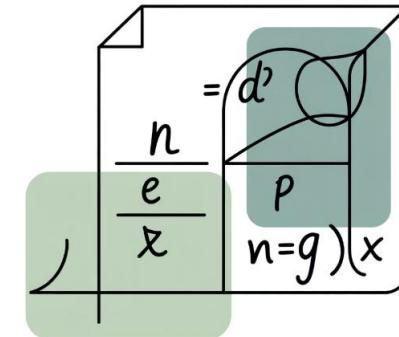
Rigor Formal e Matemático

Este tipo de trabalho foca na **construção de teorias** baseadas em definições formais e demonstrações matemáticas rigorosas. É comum em áreas matematicamente densas como teoria da complexidade, criptografia e otimização.

Elementos essenciais:

- **Formalização precisa:** Definições matemáticas claras de conceitos e estruturas
- **Axiomas e premissas:** Base lógica explícita do raciocínio
- **Demonstração formal:** Sequência lógica de passos que provam a validade
- **Corolários e implicações:** Consequências da teoria desenvolvida

Atualmente, mesmo trabalhos teóricos costumam incluir **validação experimental** para demonstrar aplicabilidade prática dos resultados formais.



Formalização

Definições matemáticas precisas estabelecem a base teórica

Prova

Demonstração formal valida as propriedades propostas

Exemplo clássico: Demonstração formal da convergência e complexidade do algoritmo de Dijkstra para o problema de menor caminho em grafos.

[1] S.V. Nagaraj, 1997, Optimal binary search trees, *Theoretical Computer Science*, v. 188, n. 1, p. 1–44.

Tipo 4: Apresentação de um Artefato Computacional

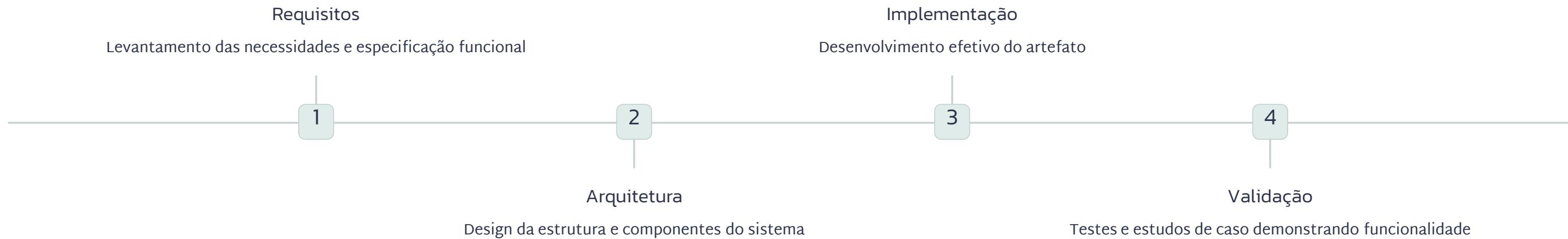
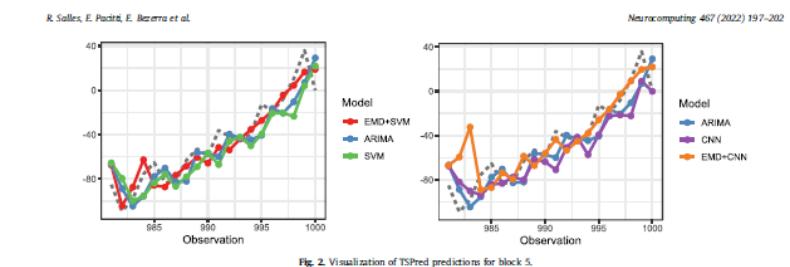
Da Teoria à Implementação

O foco está na **criação de um produto** novo — software, ferramenta, biblioteca ou framework — que resolve um problema prático. Esta abordagem é extremamente comum em Trabalhos de Conclusão de Curso (TCCs) e projetos aplicados.

Características distintivas:



- Pode não exigir rigor científico tradicional na apresentação
- Deve demonstrar impacto prático e usabilidade
- Validação através de casos de uso reais
- Documentação técnica detalhada
- Código-fonte disponibilizado (preferencialmente open-source)



Exemplo: Desenvolvimento do **TSPred**, um framework completo para predição de séries temporais não-estacionárias, incluindo pré-processamento, modelagem e visualização.

[1] R. Salles, E. Pacitti, E. Bezerra, F. Porto, and E. Ogasawara, "TSPred: A framework for nonstationary time series prediction," *Neurocomputing*, vol. 467, pp. 197–202, 2022.

Valorizando Datasets de Qualidade

Artigos de dados (*data papers*) são documentos **revisados por pares** que descrevem detalhadamente um conjunto de dados. Estes artigos valorizam o esforço significativo para coletar, organizar, limpar e documentar dados de qualidade.

Elementos fundamentais:

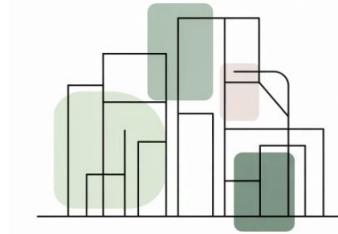
- **Processo ETL:** Extração, transformação e carga dos dados
- **Metadados completos:** Descrição detalhada de cada campo
- **Qualidade dos dados:** Estatísticas descritivas e análise de completude
- **Casos de uso potenciais:** Como a comunidade pode aproveitar os dados
- **Acesso e licenciamento:** Como obter e usar os dados legalmente

 Pipeline ETL

Documentação detalhada do processo de coleta, limpeza e transformação dos dados brutos

Baroni et al. BMC Res Notes (2020) 13:274
Page 2 of 3

Table 1 Overview of data files/data sets			
Label	Name of data file/data set	File type	Synapse ID
Data set 1	integrated_dataset	delimited text (csv)	https://www.synapse.org/#!Synapse/syn21555910 [3]
Data set 2	il_municip_Regaud	delimited text (csv)	https://www.synapse.org/#!Synapse/syn2155242 [3]
Data set 3	tb_Regaud	delimited text (csv)	https://www.synapse.org/#!Synapse/syn2155243 [3]
Data file 1	data_integration_process	R code (r)	https://www.synapse.org/#!Synapse/syn2155551 [3]
Data file 2	dictionary	delimited text (csv)	https://www.synapse.org/#!Synapse/syn2155931 [3]
Data file 3	dicionario_de_dados_slep	document (pdf)	https://www.synapse.org/#!Synapse/syn2155932 [3]
Data file 4	scheme_attribute	delimited text (csv)	https://www.synapse.org/#!Synapse/syn2156886 [3]
Data file 5	exploratory-analysis	document (pdf)	https://www.synapse.org/#!Synapse/syn2158579 [3]
Data file 6	Integrated-dataset	document (pdf)	https://www.synapse.org/#!Synapse/syn2158580 [3]



 Aplicações

Demonstração prática de como o dataset pode ser usado em pesquisas e aplicações

 Características

Estatísticas descritivas, distribuições e análise exploratória do dataset final

Processo ETL

Fluxo de preparação e validação dos dados

Uso Prático

Exemplos de aplicação do dataset em problemas reais

Exemplo de referência: O Google Open Images Dataset V4, que contém milhões de imagens anotadas para classificação, detecção de objetos e análise de relações visuais — um recurso fundamental para a comunidade de visão computacional.

Considerações Finais

Escolhendo o Caminho Certo

A pesquisa em Computação é **multifacetada e flexível**, oferecendo diversas abordagens metodológicas para diferentes tipos de problemas. Não existe uma única forma "correta" de fazer pesquisa — a escolha depende dos objetivos, contexto e natureza da contribuição.

Alinhamento com o Problema

A metodologia deve ser escolhida com base no problema específico estudado, não em preferências pessoais ou modismos acadêmicos.

Abordagens Híbridas

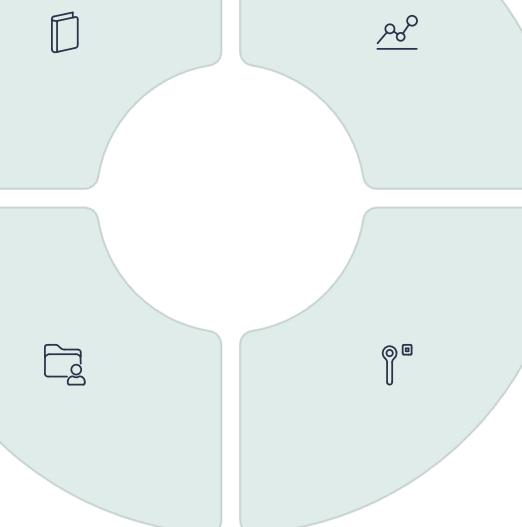
Trabalhos podem combinar diferentes metodologias: um artigo pode incluir prova formal (analítico), avaliação experimental (quantitativo) e implementação de ferramenta (tecnológico).

Equilíbrio de Focos

Apesar de ter um foco principal, pesquisas frequentemente apresentam em menor grau outras abordagens para enriquecer a contribuição.

Conhecimento Teórico

Construção de teorias e provas formais



Curadoria de Dados

Organização e disponibilização de datasets

Análise de Dados

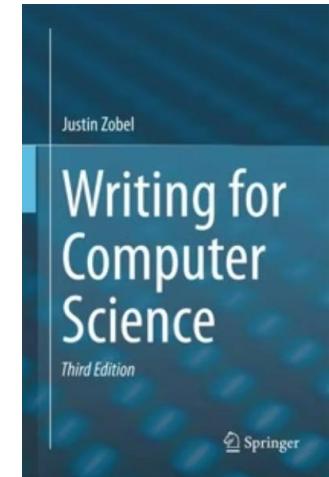
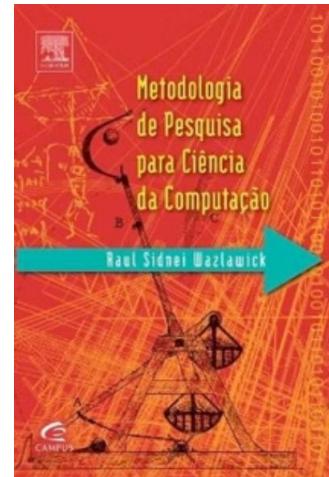
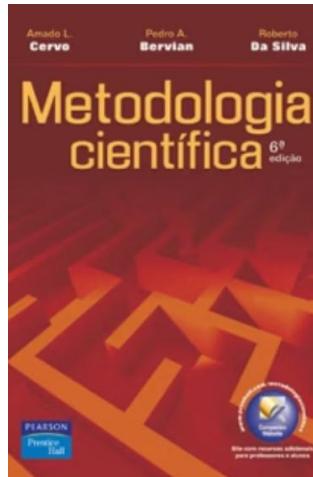
Investigação quantitativa ou qualitativa

Prototipagem

Desenvolvimento de soluções aplicadas

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.