



**CCM002**

**Metodologia de Pesquisa em Ciência da Computação**

# **Título e Estrutura de uma monografia de mestrado/doutorado**

Prof. Jesús P. Mena-Chalco  
jesus.mena@ufabc.edu.br

QS-2020

# Escrevendo a monografia (dissertação/tese)

- **Nunca viu uma dissertação/tese na vida?**

Pegue uma na biblioteca ou no site do poscomp.



- Recente
- Do seu orientador
- Veja a estrutura (divisão de capítulos)
- Linguagem
- Formas de referências
- Figuras ...
- Tabelas ...
- Abordagens ...

Tenha essa atividade '*sensorial*' desse trabalho porque ao final do curso você terá que redigir um documento similar.



# Título de uma monografia

# Títulos de monografias (sugestões gerais)

**A escolha de um bom título para o trabalho de pesquisa requer bastante paciência/adaptação:**

- Liste todas as alternativas possíveis (discuta com seus colegas).
- Qual o objeto principal?  
O objeto principal deve estar na primeira parte do título.
- O que existe no estado-da-arte?  
Sua proposta deve estar bem clara no título.
- Pesquisadores da área (e.g., computação) conseguiriam associar seu título **como sendo da mesma área?**

# Exemplos de títulos “fictícios” de monografias

- 1) **Estudo** comparativo de algoritmos genéticos.
- 2) **Um sistema inteligente** para diagnostico de segurança de websites.
- 3) **Metodologia** para processamento de imagens multi-espectrais e sua aplicação em terrenos.
- 4) **Desenvolvimento e aplicação** de algoritmos de casamento aproximado.
- 5) Algoritmos em grafos.
- 6) Sistema IHC para comunicação inteligente.

# Títulos: Sugestões para Computação?

- Considere palavras-chave de seu trabalho.
- **Evite (preferencialmente) as seguintes palavras:**
  - Metodologia ...
  - Sistema ...
  - Estudo ...
  - Proposta ...
  - Diagnóstico ...
  - Revisão sistemática ...
  - Estado da arte ...
  - Estudo de caso ...

Consulte SEMPRE  
seu (sua) orientador(a)

# Títulos: Sugestões para Computação?

- **Evite títulos genéricos (“títulos de livros”).**

- “Redes complexas em imagens”
- “Estruturas de dados em bancos de dados”
- “Práticas de programação”
- “Testes de software”

- **Evite características bem específicas:**

- “... no ano 2020”.
- “... aplicado a dados coletados no Brasil”.
- “... : estudo de caso de .... ”.

- **Evite ineditismo passageiro:**

“Um novo método .... ”.

“Um novo algoritmo .... ”.

O novo de agora não é o novo de amanhã

# Títulos: Sugestões para Computação?

- **Evite títulos muito longos:**

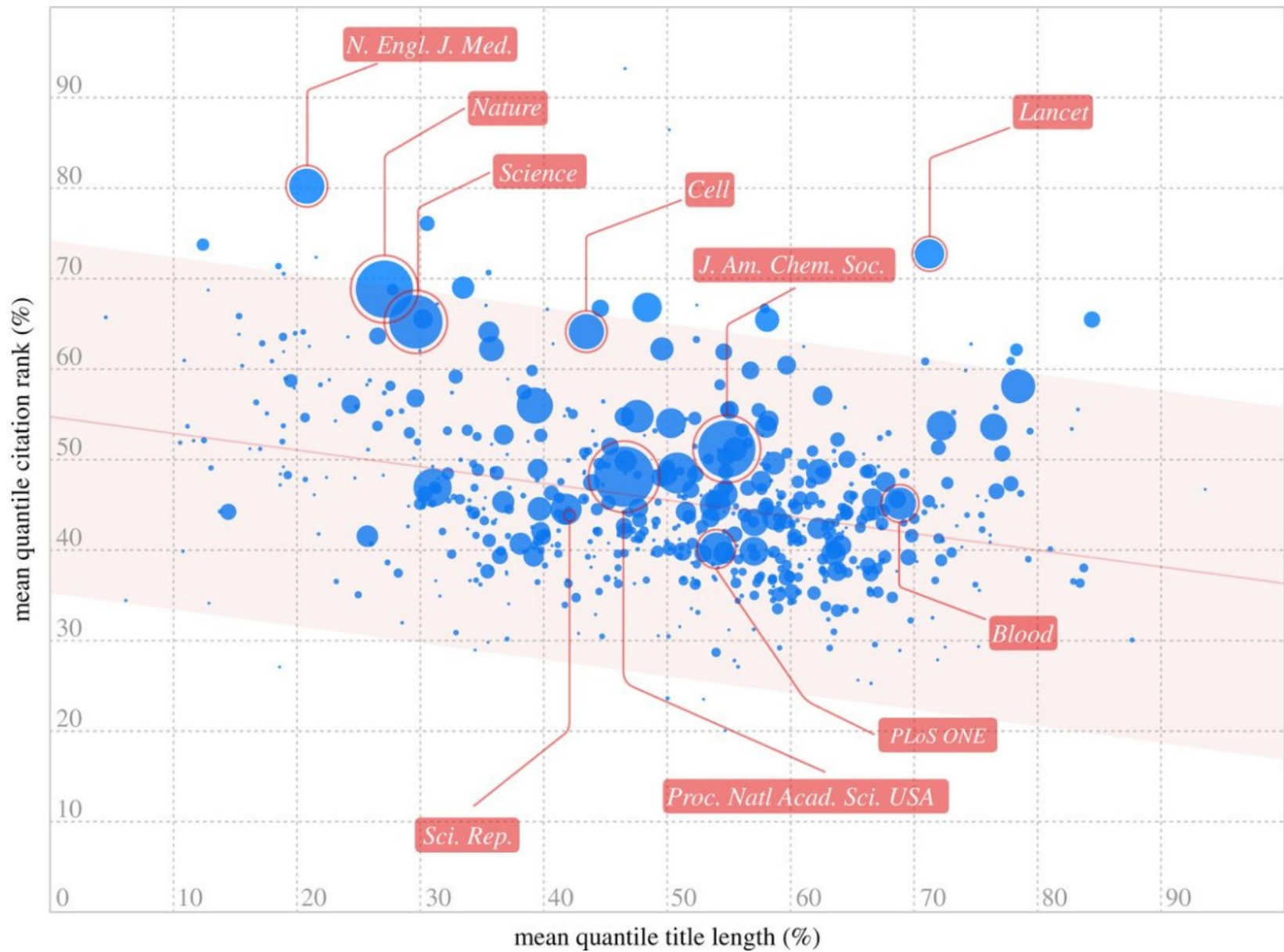
“Aplicação de Ciência da Computação no problema de pavimentação de estradas na região nordeste do Brasil.”

“Uso de bancos de dados para registro de alunos da UFABC no período de 2006-2010 e 2013-2015: Uma abordagem quantitativa.”



# Títulos em publicações

- **Claro: Nunca julgue um livro pela aparência (capa)!**
- Em 2015 Letchford *et al.* evidenciaram que a **popularidade de um artigo** pode ser estimada baseada no tamanho de seu título.
  - A brevidade parece ganhar maior atenção sobre o artigo.
- Artigos com títulos mais curtos **tendem a ser citados com maior frequência** do que aqueles com títulos mais longos.
- Amostra de 140 mil artigos (2007-2013).
- Algumas revistas limitam o número de palavras no título?



Letchford, A., Moat, H.S. and Preis, T., 2015. **The advantage of short paper titles.** *Royal Society open science*, 2(8), p.150266.

# Na Science 2010

- “The role of particle morphology in interfacial energy transfer in CDSE/CDS heterostructure nanocrystals”  
**28 citações**
- “Insects betray themselves in nature to predators by rapid isomerization of green leaf volatiles,”  
**67 citações**
- “Quantum walks of correlated photons”  
**253 citações**
- “A draft sequence of the neandertal genome”  
**700 citações**

# Títulos de monografias (sugestões gerais)

*Para pensar: Teses/dissertações com título curto são mais fáceis de ler e compreender.*

- Fique tranquilo:  
**Não há limite de número de palavras.**
- O fenômeno anterior deve ainda ser melhor discutido, **sempre teremos exceções**. Seu trabalho pode ser uma exceção, mas vale a pena analisar com calma.
- **Evite palavras redundantes:**  
Verifique se no seu título duas palavras não representam a mesma coisa.

140 Evolução da infraestrutura embarcada do projeto VERO considerando integração e migração de arcabouços de software e restrições de Tempo Real

139 Estudo Experimental do Aprovisionamento Dinâmico de Caminhos ópticos com Capacidade de Restauração Considerando Restrições na Camada Física

133 Integração, visualização e análise de informações eleitorais usando bancos de dados analíticos e fontes heterogêneas de grande volume

123 Agendamento e reserva futura de caminhos ópticos em ambientes de Lambda Grid por meio de otimização por colônia de formigas

113 Seleção de características em inferência de redes de interação gênica a partir de conjuntos reduzidos de amostras

110 Uma ontologia para interoperabilidade entre padrões de descrição de dados em biodiversidade ABCD e Darwin Core

100 Ajuste de parâmetros em algoritmos de aprendizado de máquina utilizando transferência de aprendizado

95 Inferência de redes de regulação gênica usando algoritmo de busca exaustiva em clusters de GPUs

92 Anotação Automática de Imagens Médicas Bidimensionais por meio de Classificação Multirrotulo

91 Técnicas e Aplicações de Superfícies Implícitas a partir de Dados Hermitianos Generalizados

90 Implementações de Algoritmos Paralelos da Subsequência máxima e da Submatriz Máxima em GPU

90 Roteamento ótico em redes GMPLS multidomínio por meio da otimização de colônia de formigas

88 Algoritmos paralelos em GPUS para problemas de programação quadrática binária irrestrita

87 Aplicação do protocolo SPDY para aplicativos de monitoramento sobre redes de IP público

87 Processamento de imagens HDR utilizando filtros não lineares e Decomposição Multiescala

86 Método para modelagem de processos de negócios na engenharia de requisitos de software

80 Problema da Cobertura por Caminhos com k-Terminais-Fixos em Grafos de Intervalos

79 Caracterização de grafos de genealogia acêmica por meio de métricas topológicas

78 Análise de Técnicas de Decomposição em Algoritmos de Estimação de Distribuição

78 Projeto de operadores de imagens binárias usando combinação de classificadores

77 Uma biblioteca para desenvolvimento de aplicações CUDA em aglomerados de GPUS

75 Processamento e Estilização de Dados RGB-Z em Tempo Real: uma solução móvel

73 Exploração de Novas Políticas de Construção de Métodos de Acesso Métricos

69 Análise de Funções Booleanas e Engenharia Reversa em Jogos de Minoria

67 Reconstrução Tridimensional do Corpo Humano a partir de Dados RGB-D

67 Uma abordagem arquitetural para a orquestração dinâmica de serviços

64 Método de engenharia de requisitos baseado em BPMN e caso de uso

64 Um Algoritmo Paralelo para Ciclos Hamiltonianos em Grafos Kneser

58 Métodos para Problemas de Seleção de Cadeias de Caracteres

54 Ciclos Hamiltonianos em produtos cartesianos de grafos

54 Sistema de Detecção Remota de Sonolência em motoristas

53 Um Algoritmo para o Problema do Isomorfismo de Grafos

51 Captura e Aplicação de Movimentos em Personagens 3D

50 Algoritmos genéticos em inferência de rede gênicas

59 resultados para **blockchain**

Exibindo 1-20 de 59



Refinar meus resultados

Tipo:

3 opções

- Mestrado (Dissertação) 37
- Mestrado Profissional 14

Ano:

4 opções

- 2019 35
- 2018 20
- 2016 2
- 2017 2

1. CALDAS, HELDER LUIZ PALMIERI. **Estudo da Dinamicidade do Sistema Bitcoin'** 07/03/2016 72 f. Mestrado em Ciência da Computação Instituição de Ensino: UNIVERSIDADE FEDERAL DE JUIZ DE FORA, Juiz de Fora Biblioteca Depositária: Universidade Federal de Juiz de Fora [Detalhes](#)
2. REBELLO, GABRIEL ANTONIO FONTES. **CORRENTES DE BLOCOS EM REDES VIRTUALIZADAS: PROTOCOLOS DE CONSENSO E FATIAMENTO SEGURO DA REDE'** 21/10/2019 100 f. Mestrado em ENGENHARIA ELÉTRICA Instituição de Ensino: UNIVERSIDADE FEDERAL DO RIO DE JANEIRO, Rio de Janeiro Biblioteca Depositária: Do CT - UFRJ [Detalhes](#)
3. ALMEIDA, SAMANTHA KELLY SOARES DE. **UMA ABORDAGEM PARA O DESENVOLVIMENTO DE SOFTWARES QUE UTILIZAM BLOCKCHAIN '** 22/12/2017 122 f. Mestrado em INFORMÁTICA APLICADA Instituição de Ensino: UNIVERSIDADE DE FORTALEZA, Fortaleza Biblioteca Depositária: Biblioteca Central da UNIFOR [Detalhes](#)
4. WALDRICH, CAMILA LIBERATO DE SOUSA. **A SUSTENTABILIDADE DA ATIVIDADE NOTARIAL: UMA ANÁLISE SOBRE A EVOLUÇÃO DA ATIVIDADE DOS NOTÁRIOS À LUZ DAS MUDANÇAS PARADIGMÁTICAS.'** 11/04/2018 118 f. Mestrado em CIÊNCIA JURÍDICA Instituição de Ensino: UNIVERSIDADE DO VALE DO ITAJAÍ, Itajaí Biblioteca Depositária: Biblioteca Central UNIVALI [Detalhes](#)
5. ALMEIDA, RAMON BARENCO ACETI HERDY DE. **INVEST IN DATA: AUXILIANDO AS PESSOAS A INVESTIREM MELHOR.'** 06/02/2019 32 f. Mestrado Profissional em ADMINISTRAÇÃO DO DESENVOLVIMENTO DE NEGÓCIOS Instituição de Ensino: UNIVERSIDADE PRESBITERIANA MACKENZIE, São Paulo Biblioteca Depositária: Biblioteca George Alexander [Detalhes](#)
6. TEIDER, JOSELIO JORGE. **A regulamentação no Brasil dos Contratos Inteligentes implementados pela tecnologia Blockchain'** 20/03/2019 141 f. Mestrado em DIREITO Instituição de Ensino: PONTIFÍCIA UNIVERSIDADE CATÓLICA DO PARANÁ, Curitiba Biblioteca Depositária: undefined [Detalhes](#)
7. CARVALHO, JOSELMA DA ROCHA. **Perspectivas da utilização da tecnologia blockchain nos Correios.'** 14/12/2018 169 f. Mestrado Profissional em Governança, Tecnologia e Inovação Instituição de Ensino: UNIVERSIDADE CATÓLICA DE BRASÍLIA, Brasília Biblioteca Depositária: UCB [Detalhes](#)

# Prêmio Teses CAPES 2019



Área de avaliação	Autor	Programa	IES	Orientador	Coorientador
CIÊNCIA DA COMPUTAÇÃO	ALLAN DA SILVA PINTO	CIÊNCIA DA COMPUTAÇÃO	UNICAMP	ANDERSON DE REZENDE ROCHA	HÉLIO PEDRINI

## Menção honrosa

CIÊNCIA DA COMPUTAÇÃO	FABIANO MUNIZ BELEM	CIÊNCIAS DA COMPUTAÇÃO	UFMG	JUSSARA MARQUES DE ALMEIDA GONÇALVES	MARCOS ANDRÉ GONÇALVES
CIÊNCIA DA COMPUTAÇÃO	TROY COSTA KOHWALTER	COMPUTAÇÃO	UFF	ESTEBAN WALTER GONZALEZ CLUA	LEONARDO GRESTA PAULINO MURTA

<b>Instituição de Ensino Superior:</b>	UNIVERSIDADE ESTADUAL DE CAMPINAS
<b>Programa:</b>	CIÊNCIA DA COMPUTAÇÃO (33003017005P8)
<b>Título:</b>	Analysis of Intrinsic and Extrinsic Properties of Biometric Samples for Presentation Attack Detection
<b>Autor:</b>	ALLAN DA SILVA PINTO
<b>Tipo de Trabalho de Conclusão:</b>	TESE
<b>Data Defesa:</b>	06/09/2018
<b>Resumo:</b>	<p>Os recentes avanços nas áreas de pesquisa em biometria, forense e segurança da informação trouxeram importantes melhorias na eficácia dos sistemas de reconhecimento biométricos. No entanto, um desafio crescente em biometria é a vulnerabilidade de tais sistemas contra ataques de apresentação, nos quais os usuários impostores criam amostras sintéticas a partir das informações biométricas originais de um usuário legítimo e as mostram ao sensor de aquisição procurando se autenticar como usuários válidos. Dependendo da característica biométrica usada para a autenticação dos usuários, os tipos de ataque variam de acordo com o tipo de material usado para construir as amostras sintéticas. Por exemplo, em biometria facial, uma tentativa de ataque é caracterizada quando um usuário impostor apresenta ao sensor de aquisição uma fotografia, um vídeo digital ou uma máscara 3D com as informações faciais de um usuário-alvo. Em sistemas de biometria baseados em íris, os ataques de apresentação podem ser realizados com fotografias impressas ou com lentes de contato contendo os padrões de íris de um usuário-alvo. Nos sistemas biométricos de impressão digital, os usuários impostores podem enganar o processo de autenticação usando réplicas dos padrões de impressão digital construídos com materiais sintéticos, como látex, massa de modelar, silicone, entre outros. Esta pesquisa tem como objetivo o desenvolvimento de soluções para detecção de ataques de apresentação considerando os sistemas biométricos faciais, de íris e de impressão digital. As linhas de investigação apresentadas nesta tese incluem o desenvolvimento de representações baseadas nas informações espaciais, temporais e espectrais da assinatura de ruído; em propriedades intrínsecas das amostras biométricas (por exemplo, albedo, reflectância e mapas de profundidade) e em técnicas de aprendizagem supervisionada de características. Os principais resultados e contribuições apresentadas nesta tese incluem: o desenvolvimento de novas abordagens para modelagem e análise de propriedades extrínsecas das amostras biométricas relacionadas aos artefatos que são adicionados durante a fabricação das amostras sintéticas e sua captura pelo sensor de aquisição; a investigação de uma abordagem baseada na análise de propriedades intrínsecas das superfícies faciais estimadas a partir da informação de sombras presentes em sua superfície como, por exemplo, a estimação de mapas de albedo, reflectância e de profundidade; e, por fim, a investigação de diferentes abordagens baseadas em redes neurais convolucionais para o aprendizado automático de características relacionadas ao nosso problema incluindo o projeto de eficientes redes neurais com arquiteturas rasas capazes de aprender características relacionadas ao nosso problema a partir dos pequenos conjuntos de dados disponíveis para o desenvolvimento e a avaliação de soluções para a detecção de ataques de apresentação.</p>
<b>Palavras-Chave:</b>	Sistemas Biométricos;Detecção de Ataque de Apresentação;Ritmos Visuais;UVAD;Análise Espectral;Dicionários Visuais;Características Visuais Temporais-Espectrais;Reconstrução de superfície;CNN;Otimização de hiperparâmetros;Aprendizado Profundo
<b>Abstract:</b>	Recent advances in biometrics, information forensics, and security have improved the recognition effectiveness of biometric systems. However, an ever-growing challenge is the vulnerability of such systems against presentation attacks, in which impostor users create synthetic samples from the original biometric information of a legitimate user and show them to the acquisition sensor seeking to authenticate themselves as legitimate users. Depending on the trait used by the biometric authentication, the attack types vary with the type of material used to build the synthetic samples. For instance, in facial biometric systems, an attempted attack is characterized by the type of material the impostor uses such as a photograph, a digital video, or a 3D mask with the facial information of a target user. In iris-based biometrics, presentation attacks can be accomplished with printout photographs or with contact lenses containing the iris patterns of a target user. In fingerprint biometric systems, impostor users can deceive the authentication process using replicas of the fingerprint patterns built with synthetic materials such as latex, play-doh, silicone, among others. This research aims at developing presentation attack detection (PAD) solutions whose objective is to detect attempted attacks considering these different attack types, in each modality. The lines of investigation presented in this thesis aims to devise and develop representations based on spatial, temporal and spectral information from noise signature, intrinsic properties of the biometric data (for instance, albedo, reflectance, and depth maps), and supervised feature learning techniques, taking into account different testing scenarios including cross-sensor, intra-, and inter-dataset scenarios. The main findings and contributions presented in this thesis include: the development of novel approaches to modeling and analysis of extrinsic properties of biometric samples related to the artifacts added during the manufacturing of the synthetic samples and their capture by the acquisition sensor; the investigation of an approach based on the analysis of intrinsic properties of facial surfaces, estimated from the information of shadows present on their surface; and the investigation of different approaches to automatically learning representations related to our problem, which includes the design of efficient neural networks with shallow architectures capable of learning characteristics related to our problem from the small sets of data available to develop and evaluate PAD solutions.
<b>Keyword:</b>	Biometric Systems;Presentation Attack Detection;Visual Rhythm;UVAD dataset;Spectral Analysis;Visual Codebook;Time-spectral Visual Features;Surface Reconstruction;CNN;Hyperparameter Architecture Optimization;Deep Learning
<b>Volume:</b>	1
<b>Páginas:</b>	208
<b>Idioma:</b>	INGLES



<b>1</b>	<b>Introduction</b>	<b>24</b>
1.1	Basic Concepts in Biometrics . . . . .	25
1.2	Presentation Attacks in Biometrics . . . . .	25
1.3	Research Vision: New Insights for the Presentation Attack Detection Problem	28
1.3.1	Problems with the Existing Approaches . . . . .	28
1.3.2	Hypothesis Statements . . . . .	29
1.3.3	Novelties and Rationales Brought in this Thesis . . . . .	30
1.4	Key Contributions . . . . .	32
1.5	Thesis Organization . . . . .	34
<b>2</b>	<b>Using Visual Rhythms for Detecting Video-based Facial Spoof Attacks</b>	<b>36</b>
2.1	Introduction . . . . .	37
2.2	Related Work . . . . .	40
2.2.1	Existing Databases . . . . .	41
2.2.2	Motion Analysis and Clues of the Scene . . . . .	42
2.2.3	Texture and Frequency Analysis . . . . .	43
2.2.4	Other Approaches . . . . .	44
2.2.5	Problems with the Existing Approaches . . . . .	45
2.3	Proposed Method . . . . .	46
2.3.1	Calculation of the Residual Noise Videos . . . . .	47
2.3.2	Calculation of the Fourier Spectrum Videos . . . . .	47
2.3.3	Calculation of the Visual Rhythms . . . . .	47
2.3.4	Feature Extraction . . . . .	49
2.3.5	Learning . . . . .	50
2.4	Database Creation . . . . .	51
2.5	Experimental Results . . . . .	52
2.5.1	Protocols for the UVAD Database . . . . .	52
2.5.2	Parameters for the Filtering Process, Visual Rhythm Analysis and Classification . . . . .	53
2.5.3	Experiment I: Finding the Best Configuration . . . . .	54
2.5.4	Experiment II: Influence of the Biometric Sensors . . . . .	55
2.5.5	Experiment III: Influence of the Display Devices . . . . .	56
2.5.6	Experiment IV: Comparison to a State-of-the-Art Method for Photo-Based Spoofing Attack Detection . . . . .	57
2.5.7	Experiment V: Evaluation of the Method in the Replay-Attack Database . . . . .	58
2.6	Conclusions and Future Work . . . . .	59

<b>3</b>	<b>Face Spoofing Detection Through Visual Codebooks of Spectral Temporal Cubes</b>	<b>61</b>
3.1	Introduction . . . . .	62
3.2	Related Work . . . . .	65
3.2.1	Frequency-based approaches . . . . .	65
3.2.2	Texture-based approaches . . . . .	66
3.2.3	Motion-based approaches . . . . .	67
3.3	Proposed Method . . . . .	67
3.3.1	Low-Level Descriptor Extraction . . . . .	68
3.3.2	Mid-Level Descriptor Extraction . . . . .	71
3.3.3	Classification . . . . .	73
3.4	Experiments and Results . . . . .	73
3.4.1	Datasets . . . . .	73
3.4.2	Experimental Protocol . . . . .	74
3.4.3	Method Parameterization . . . . .	76
3.4.4	Experimental Design and Analysis . . . . .	76
3.4.5	Summary After Analyzing Different Factors and Levels . . . . .	79
3.4.6	Results . . . . .	80
3.5	Conclusions and Future Work . . . . .	86
<b>4</b>	<b>Deep Representations for Iris, Face, and Fingerprint Spoofing Detection</b>	<b>89</b>
4.1	Introduction . . . . .	90
4.2	Related Work . . . . .	93
4.2.1	Iris Spoofing . . . . .	93
4.2.2	Face Spoofing . . . . .	94
4.2.3	Fingerprint Spoofing . . . . .	95
4.2.4	Multi-modalities . . . . .	95
4.3	Benchmarks . . . . .	96
4.3.1	Iris Spoofing Benchmarks . . . . .	96
4.3.2	Video-based Face Spoofing Benchmarks . . . . .	96
4.3.3	Fingerprint Spoofing Benchmarks . . . . .	98
4.3.4	Remark . . . . .	98
4.4	Methodology . . . . .	98
4.4.1	Architecture Optimization (AO) . . . . .	98
4.4.2	Filter Optimization (FO) . . . . .	101
4.4.3	Elementary Preprocessing . . . . .	103
4.4.4	Evaluation Protocol . . . . .	104
4.4.5	Implementation . . . . .	105
4.5	Experiments and Results . . . . .	105
4.5.1	Architecture Optimization (AO) . . . . .	107
4.5.2	Filter Optimization (FO) . . . . .	107
4.5.3	Interplay between AO and FO . . . . .	110
4.5.4	Runtime . . . . .	111
4.5.5	Visual Assessment . . . . .	112
4.6	Conclusions and Future Work . . . . .	113

<b>5 Counteracting Presentation Attacks in Face, Fingerprint, and Iris Recognition</b>	<b>115</b>
5.1 Introduction	116
5.2 Related Work	118
5.2.1 Face Presentation Attack Detection	118
5.2.2 Fingerprint Presentation Attack Detection	121
5.2.3 Iris Presentation Attack Detection	123
5.2.4 Unified Frameworks to Presentation Attack Detection	125
5.3 Methodology	126
5.3.1 Network Architecture	126
5.3.2 Training and Testing	126
5.3.3 Memory Footprint	128
5.4 Metrics and Datasets	128
5.4.1 Video-based Face Spoofing Benchmarks	128
5.4.2 Fingerprint Spoofing Benchmarks	129
5.4.3 Iris Spoofing Benchmarks	130
5.4.4 Error Metrics	131
5.5 Results	132
5.5.1 Face	132
5.5.2 Fingerprints	134
5.5.3 Iris	136
5.6 Conclusions	140
<b>6 Leveraging Shape, Reflectance and Albedo from Shading for Face Presentation Attack Detection</b>	<b>143</b>
6.1 Introduction	144
6.1.1 Optical Properties of the Light and Rationale of Proposed Approach	145
6.1.2 Contributions and Organization	146
6.2 Related Work	147
6.3 Proposed Method	149
6.3.1 Surface Reconstruction: Recovering the Depth, Reflectance and Albedo maps	149
6.3.2 Convolutional Neural Network for Learning Intrinsic Surface Properties	151
6.4 Experimental Results	152
6.4.1 Datasets	152
6.4.2 Experimental Protocols	153
6.4.3 Experimental Setup	154
6.4.4 Evaluation of the Proposed CNN Architecture	154
6.4.5 How to Feed the Proposed CNN Network with the Different Maps?	155
6.4.6 Intra-dataset Evaluation Protocol	156
6.4.7 Inter-dataset Evaluation Protocol	159
6.4.8 Comparison with State-of-the-Art Methods	159
6.4.9 Visual Assessment	161
6.5 Conclusion and Future Work	162
<b>7 Conclusions and Future Work</b>	<b>164</b>
7.1 Final Remarks	164
7.2 Directions for Future Work	165
7.3 Other Applications to Algorithms Presented in this Thesis	166
7.3.1 Detection of (Illegal) Copyrighted Video Recapture	166
7.3.2 Image Tampering Detection	167
7.4 Publications During this Doctoral Research	167
7.5 Source Code Available Along with this Thesis	171
<b>Bibliography</b>	<b>172</b>
<b>A Ethics Committee Approval</b>	<b>191</b>
<b>B Convolutional Network Operations</b>	<b>197</b>
<b>C Copyright Permissions</b>	<b>199</b>

# Prêmio Teses CAPES 2020

[https://www.gov.br/capes/pt-br/centrais-de-conteudo/resultados-dos-editais/01102020\\_EDITAL102020\\_Resultado.pdf](https://www.gov.br/capes/pt-br/centrais-de-conteudo/resultados-dos-editais/01102020_EDITAL102020_Resultado.pdf)

Área de avaliação	Autor	Título da Tese	Programa	IES	Orientador	Coorientador
CIÊNCIA DA COMPUTAÇÃO	HENRIQUE FERRAZ DE ARRUDA	Multi-scale analysis of languages and knowledge through complex networks	CIÊNCIAS DA COMPUTAÇÃO E MATEMÁTICA COMPUTACIONAL	USP/SC	LUCIANO DA FONTOURA COSTA	DIEGO RAPHAEL AMANCIO

## Menção honrosa

CIÊNCIA DA COMPUTAÇÃO	GUSTAVO FREITAS SANCHEZ	EXPLORATION OF ALGORITHMS AND IMPLEMENTATIONS FOR EFFICIENT 3D-HEVC DEPTH MAP ENCODING	CIÊNCIA DA COMPUTAÇÃO	PUC/RS	CÉSAR AUGUSTO MISSIO MARCON	LUCIANO VOLCAN AGOSTINI
CIÊNCIA DA COMPUTAÇÃO	MICHEL MELO DA SILVA	HYPERLAPSE SEMÂNTICO PARA VÍDEOS EM PRIMEIRA PESSOA: UMA ABORDAGEM MULTI-IMPORTÂNCIA BASEADA EM CODIFICAÇÃO ESPARSA	CIÊNCIAS DA COMPUTAÇÃO	UFMG	ERICKSON RANGEL DO NASCIMENTO	MARIO FERNANDO MONTENEGRO CAMPOS

# Prêmio Teses CAPES 2020

[https://teses.usp.br/teses/disponiveis/55/55134/tde-26042019-105207/publico/HenriqueFerrazdeArruda\\_revisada.pdf](https://teses.usp.br/teses/disponiveis/55/55134/tde-26042019-105207/publico/HenriqueFerrazdeArruda_revisada.pdf)

1	INTRODUCTION . . . . .	37	3	NETWORK-BASED APPROACHES TO TEXT MINING . . . . .	79
1.1	Objectives . . . . .	41	3.1	Databases . . . . .	79
1.2	Contributions . . . . .	42	3.1.1	<i>Brown Database</i> . . . . .	79
1.3	Thesis organization . . . . .	45	3.1.2	<i>Gutenberg</i> . . . . .	80
2	BASIC CONCEPTS . . . . .	47	3.1.3	<i>Holy Bible dataset</i> . . . . .	81
2.1	Complex network definitions . . . . .	47	3.1.4	<i>Voynich Manuscript</i> . . . . .	81
2.2	Characterization of complex networks . . . . .	48	3.1.5	<i>Artificial texts</i> . . . . .	81
2.2.1	<i>Local measurements</i> . . . . .	48	3.2	Classification between imaginative and informative texts . . . . .	82
2.2.2	<i>Global measurements</i> . . . . .	52	3.2.1	<i>From texts to co-occurrence networks</i> . . . . .	83
2.2.3	<i>Network communities</i> . . . . .	57	3.2.2	<i>Pattern recognition and evaluation</i> . . . . .	84
2.3	Network models . . . . .	58	3.2.3	<i>Results and discussion</i> . . . . .	84
2.3.1	<i>Erdős-Rényi model</i> . . . . .	58	3.2.4	<i>Conclusions</i> . . . . .	89
2.3.2	<i>Watts-Strogatz model</i> . . . . .	58	3.3	Network-based subject detection . . . . .	90
2.3.3	<i>Barabási-Albert model</i> . . . . .	59	3.3.1	<i>Network approaches</i> . . . . .	90
2.3.4	<i>Network model of communities</i> . . . . .	61	3.3.2	<i>Detection of subjects via community detection in networks</i> . . . . .	92
2.3.5	<i>Waxman model</i> . . . . .	62	3.3.3	<i>Modularity analysis</i> . . . . .	93
2.3.6	<i>Random geometric graph</i> . . . . .	63	3.3.4	<i>Results and discussion</i> . . . . .	95
2.3.7	<i>Knitted</i> . . . . .	64	3.3.5	<i>Conclusions</i> . . . . .	97
2.4	Network visualization . . . . .	65	3.4	Paragraph-based networks . . . . .	100
2.5	Complex networks dynamics . . . . .	66	3.4.1	<i>Paragraph-based network approach</i> . . . . .	100
2.6	Characteristics of natural languages . . . . .	67	3.4.2	<i>Network variations</i> . . . . .	101
2.6.1	<i>Text pre-processing</i> . . . . .	68	3.4.3	<i>Informativeness analysis</i> . . . . .	102
2.6.2	<i>Bag-of-words</i> . . . . .	69	3.4.4	<i>Dependency with language and semantics</i> . . . . .	103
2.6.2.1	<i>Tf-idf</i> . . . . .	70	3.4.5	<i>Results and discussion</i> . . . . .	103
2.6.2.2	<i>Cosine similarity</i> . . . . .	71	3.4.5.1	<i>Informativeness analysis</i> . . . . .	104
2.6.2.3	<i>Latent Semantic Analysis</i> . . . . .	71	3.4.5.2	<i>Dependency with language and semantics</i> . . . . .	108
2.6.3	<i>Bag-of-character</i> . . . . .	72	3.4.5.3	<i>Classification tests</i> . . . . .	108
2.7	Complex network-based models . . . . .	72	3.4.5.3.1	Discriminating between real and shuffled texts . . . . .	108
2.7.1	<i>Word networks</i> . . . . .	72	3.4.5.3.2	Case Example: Voynich manuscript . . . . .	111
2.7.2	<i>Topic segmentation</i> . . . . .	73	3.4.6	<i>Conclusions</i> . . . . .	112
2.7.3	<i>Applications of complex networks to text mining problems</i> . . . . .	75	3.5	Mesoscopic analysis of texts . . . . .	114
			3.5.1	<i>Mesoscopic network approach</i> . . . . .	114
			3.5.2	<i>Classification between real and shuffled texts</i> . . . . .	115
			3.5.3	<i>Image-based features</i> . . . . .	116
			3.5.4	<i>Results and discussion</i> . . . . .	117
			3.5.4.1	Case study: Mesoscopic analysis of "Alice's adventures in wonderland" . . . . .	118
			3.5.4.2	Discriminating real from shuffled texts . . . . .	120
			3.5.4.3	Use case: Authorship attribution . . . . .	127
			3.5.5	<i>Conclusions</i> . . . . .	130

# Prêmio Teses CAPES 2020

[https://teses.usp.br/teses/disponiveis/55/55134/tde-26042019-105207/publico/HenriqueFerrazdeArruda\\_revisada.pdf](https://teses.usp.br/teses/disponiveis/55/55134/tde-26042019-105207/publico/HenriqueFerrazdeArruda_revisada.pdf)

4	<b>ANALYSIS OF LANGUAGE AND KNOWLEDGE VIA NETWORK DYNAMICS</b> . . . . .	133
4.1	<b>Datasets</b> . . . . .	133
4.1.1	<i>APS dataset</i> . . . . .	134
4.1.2	<i>Knowledge representation</i> . . . . .	134
4.2	<b>Integration between theoretical and applied physics</b> . . . . .	134
4.2.1	<i>Methods</i> . . . . .	134
4.2.2	<i>Results and discussion</i> . . . . .	136
4.2.2.1	<i>Case Study #1 - Phys. Rev. Applied</i> . . . . .	136
4.2.2.2	<i>Case Study #2 - Phys. Rev. B</i> . . . . .	137
4.2.3	<i>Conclusions</i> . . . . .	138
4.3	<b>Simulating knowledge acquisition</b> . . . . .	139
4.3.1	<i>Problem definition</i> . . . . .	140
4.3.2	<i>Results</i> . . . . .	141
4.3.2.1	<i>Global analysis</i> . . . . .	142
4.3.2.2	<i>Analysis by regions</i> . . . . .	144
4.3.3	<i>Conclusions</i> . . . . .	148
4.4	<b>Intelligent complex networks</b> . . . . .	150
4.4.1	<i>Dynamics definitions</i> . . . . .	150
4.4.2	<i>Results</i> . . . . .	151
4.4.2.1	<i>Case 1: Infinite Buffer Size</i> . . . . .	152
4.4.2.2	<i>Case 2: Limited Buffer Size</i> . . . . .	154
4.4.3	<i>Conclusions</i> . . . . .	156
4.5	<b>Connecting network science and information theory</b> . . . . .	160
4.5.1	<i>Problem definition</i> . . . . .	160
4.5.2	<i>Random walk dynamics</i> . . . . .	162
4.5.3	<i>Analysis of network compression</i> . . . . .	162
4.5.4	<i>Results and discussion</i> . . . . .	163
4.5.5	<i>Use case: language compression</i> . . . . .	165
4.5.6	<i>Conclusions</i> . . . . .	168
5	<b>CONCLUSIONS AND FUTURE WORKS</b> . . . . .	171
	<b>BIBLIOGRAPHY</b> . . . . .	177
	<b>APPENDIX A LIST OF PUBLICATIONS</b> . . . . .	199
	<b>APPENDIX B LIST OF STOP WORDS</b> . . . . .	201
	<b>APPENDIX C MACHINE LEARNING</b> . . . . .	203
	<b>C.1 Standardization</b> . . . . .	203
	<b>C.2 Attribute selection</b> . . . . .	203
	<b>C.2.1 Information gain</b> . . . . .	204
	<b>C.2.2 SVM attribute selection</b> . . . . .	204
	<b>C.3 Classification</b> . . . . .	204
	<b>C.3.1 Naive Bayes</b> . . . . .	204
	<b>C.3.2 KNN</b> . . . . .	205
	<b>C.3.3 CART</b> . . . . .	205
	<b>C.3.4 SVM</b> . . . . .	206
	<b>C.4 Data clustering</b> . . . . .	206
	<b>C.4.1 K-means</b> . . . . .	206
	<b>C.4.2 Expectation maximization</b> . . . . .	206
	<b>C.5 Quantification of relevance of the characteristics</b> . . . . .	207
	<b>APPENDIX D IMAGE ANALYSIS</b> . . . . .	209
	<b>APPENDIX E INFORMATION THEORY</b> . . . . .	211
	<b>E.1 Entropy</b> . . . . .	211
	<b>E.2 Information gain</b> . . . . .	212
	<b>E.3 Huffman</b> . . . . .	212
	<b>APPENDIX F LANGUAGE COMPARISON: SYLLABLES-BASED APPROACH</b> . . . . .	213
	<b>F.1 Syllable dataset</b> . . . . .	213
	<b>F.2 Language model</b> . . . . .	214
	<b>F.2.1 Network model</b> . . . . .	215
	<b>F.2.2 Results</b> . . . . .	215



# **Estrutura de uma monografia**

## Documento para qualificação

## Documento para defesa

Pré-textuais

### Título

Resumo (palavras-chave)

Abstract (keywords)

### Título

Resumo (palavras-chave)

Abstract (keywords)

Sumário

Lista de figuras, tabelas, abreviaturas

Textuais

1) Introdução.

2) Trabalhos relacionados.  
(Referencial teórico)

3) **Objetivos** e relevância.

4) Proposta (solução).

5) **Cronograma.**

6) **Considerações finais.**

1) Introdução.

2) Trabalhos relacionados.  
(Referencial teórico)

3) **Objetivos** e relevância.

4) Proposta (solução).

5) **Experimentos (resultados).**

6) **Conclusão.**

Pós-textuais

Referências

Referências

Anexos

# Escrevendo a dissertação/tese

## Dividir para conquistar

- Assim que definir o tópico do trabalho, comece a trabalhar no documento final.
- Procure um modelo sugerido no site do Poscomp (use LaTeX).
- Comece com as páginas obrigatórias.

- 1) Capas
- 2) Catalogação
- 3) Aprovação
- 4) Dedicatória
- 5) Agradecimentos
- 6) Resumo
- 7) Abstract
- 8) Sumário
- 9) Lista de figuras
- 10) Lista de tabelas
- 11) Corpo do texto**
- 12) Apêndice
- 13) Referências



# Escrevendo a dissertação/tese

- A medida que define seu trabalho, **adicione conteúdo ao texto.**
- **Escreva rascunhos.**  
Não precisa estar bem escrito (ainda).
  - No final do trabalho será muito complicado se lembrar de tudo nos mínimos detalhes.
  - Anote as referências e faça resumos dos trabalhos que tenha lido (bibTeX completo).
  - **Cópia literal é plágio:** faça resumo de ideias relevantes em cada artigo lido.

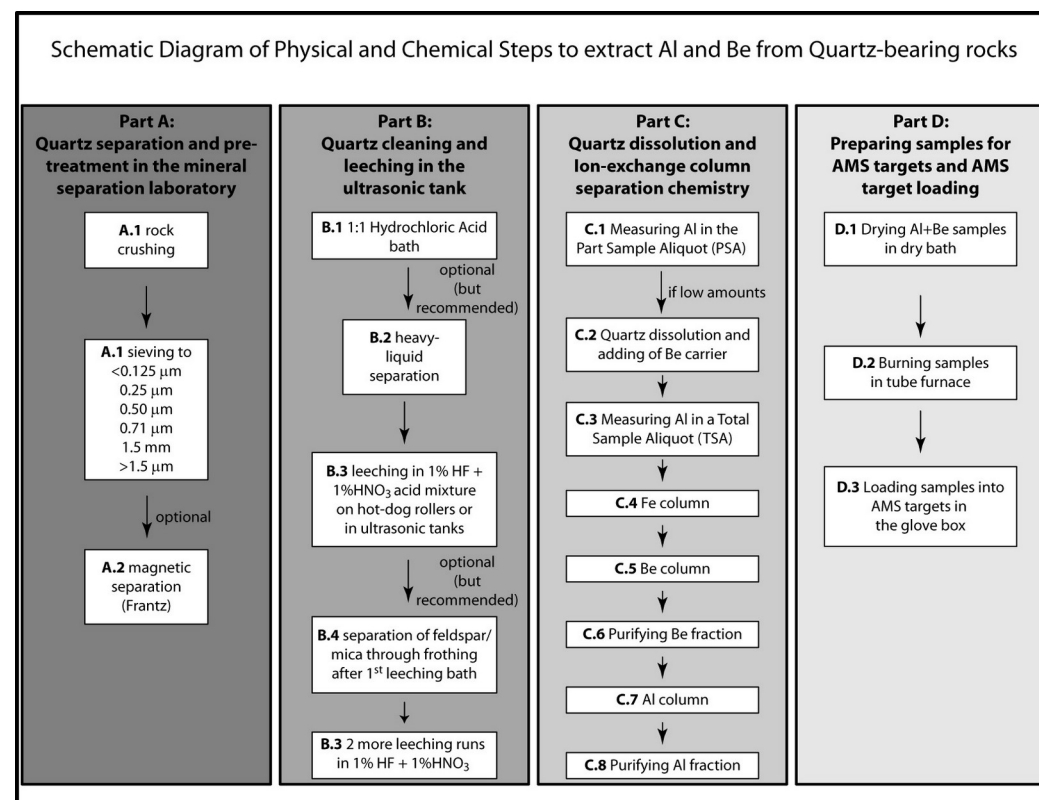
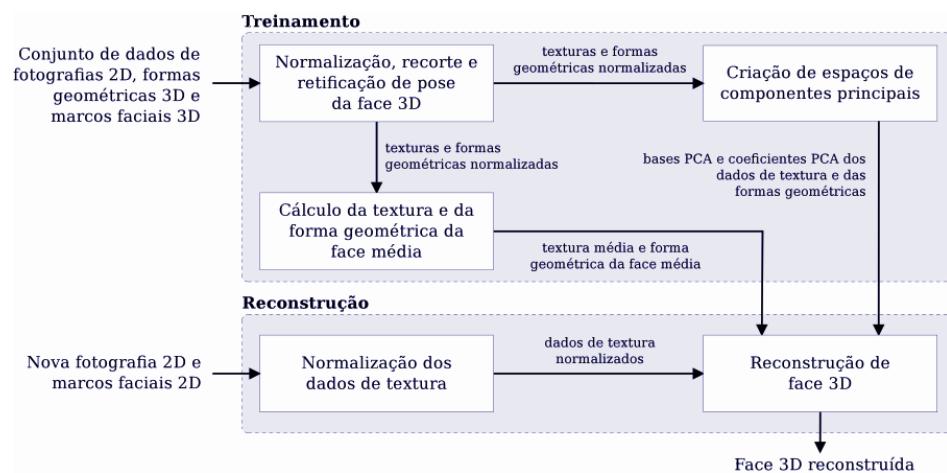
# Escrevendo a dissertação/tese

- Erro de **português é inadmissível** em dissertação/tese (manuscrito e apresentação):
  - Use as ferramentas livremente disponíveis.
    - Aspell, por exemplo.
    - Cogroo, corretor gramatical.
    - Grammarly.



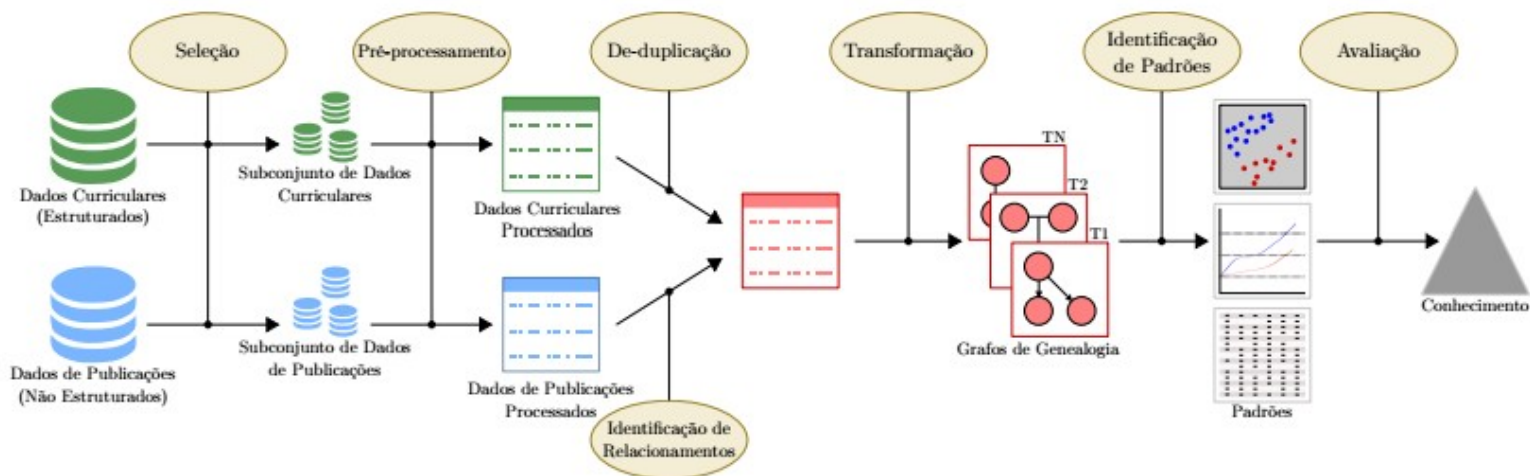
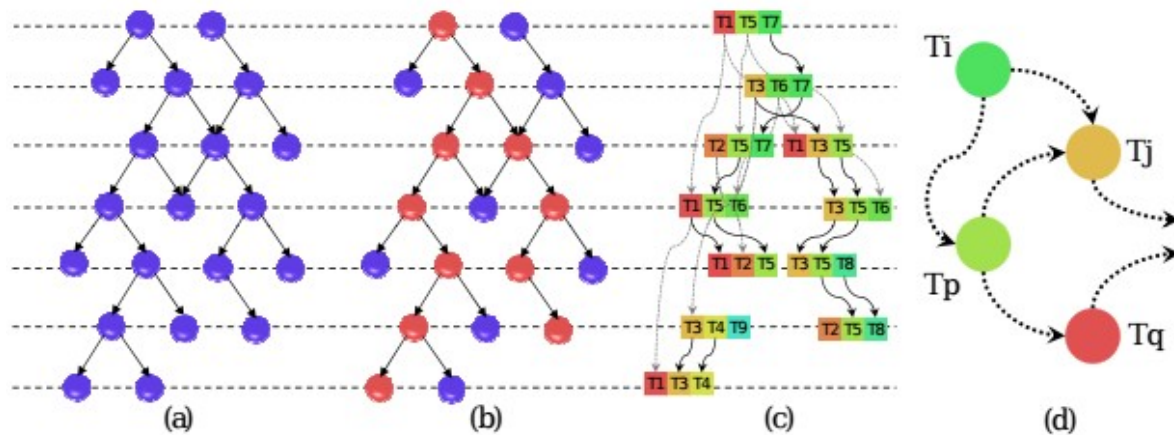
# Escrevendo a dissertação/tese

- Use de Diagramas, Figuras e tabelas.



# Escrevendo a dissertação/tese

- Use de Diagramas, Figuras e tabelas.



# Escrevendo a dissertação/tese

- Estabeleça um cronograma e tente segui-lo.

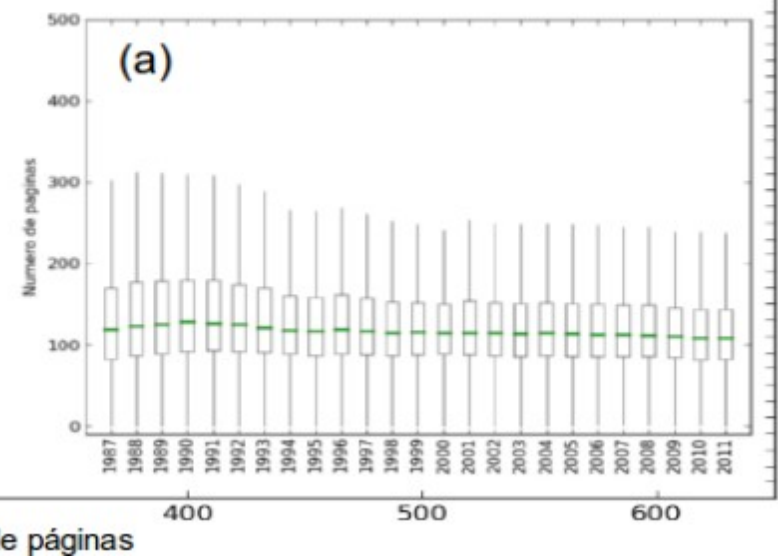
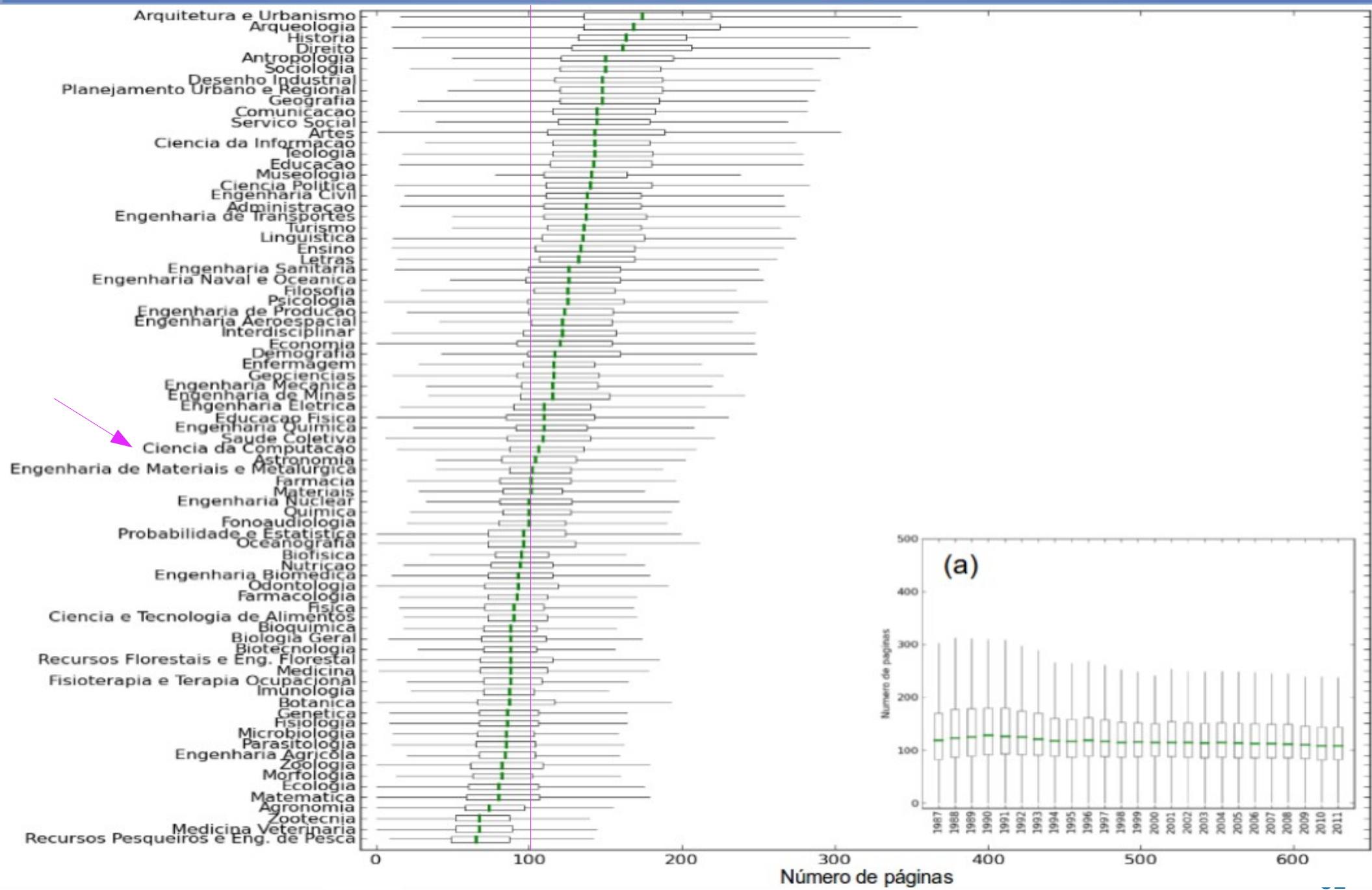
Atividade	2016			2017			2018			2019		
	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2	Q3	Q1	Q2
1. Elaboração do projeto enviado à FAPESP			o									
2. Estudo de disciplinas obrigatórias (créditos)	o	o	o	X								
3. Revisão bibliográfica	o	o	o	X	X	X	X	X	X	X	X	,
4. Prospecção dos dados e criação do grafo de genealogia		o	o	X	X							
5. Classificação e criação do subgrafo de genealogia				X	X	X						
6. Escrita de artigo científico I					X	X						
7. Representação de tópicos de atuação e criação do grafo					X	X	X					
8. Qualificação de doutorado						X	X					
9. Realização de análises experimentais						X	X	X	X			
10. Escrita de artigo científico II							X	X	X			
11. Escrita da tese									X	X		
12. Defesa da tese												X



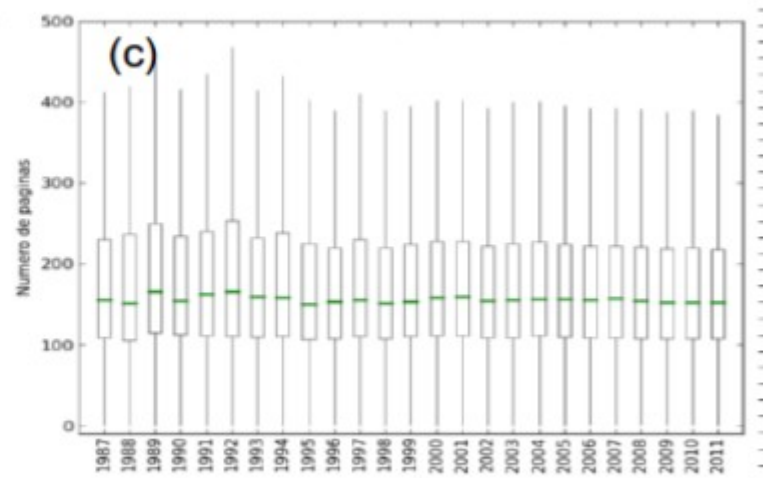
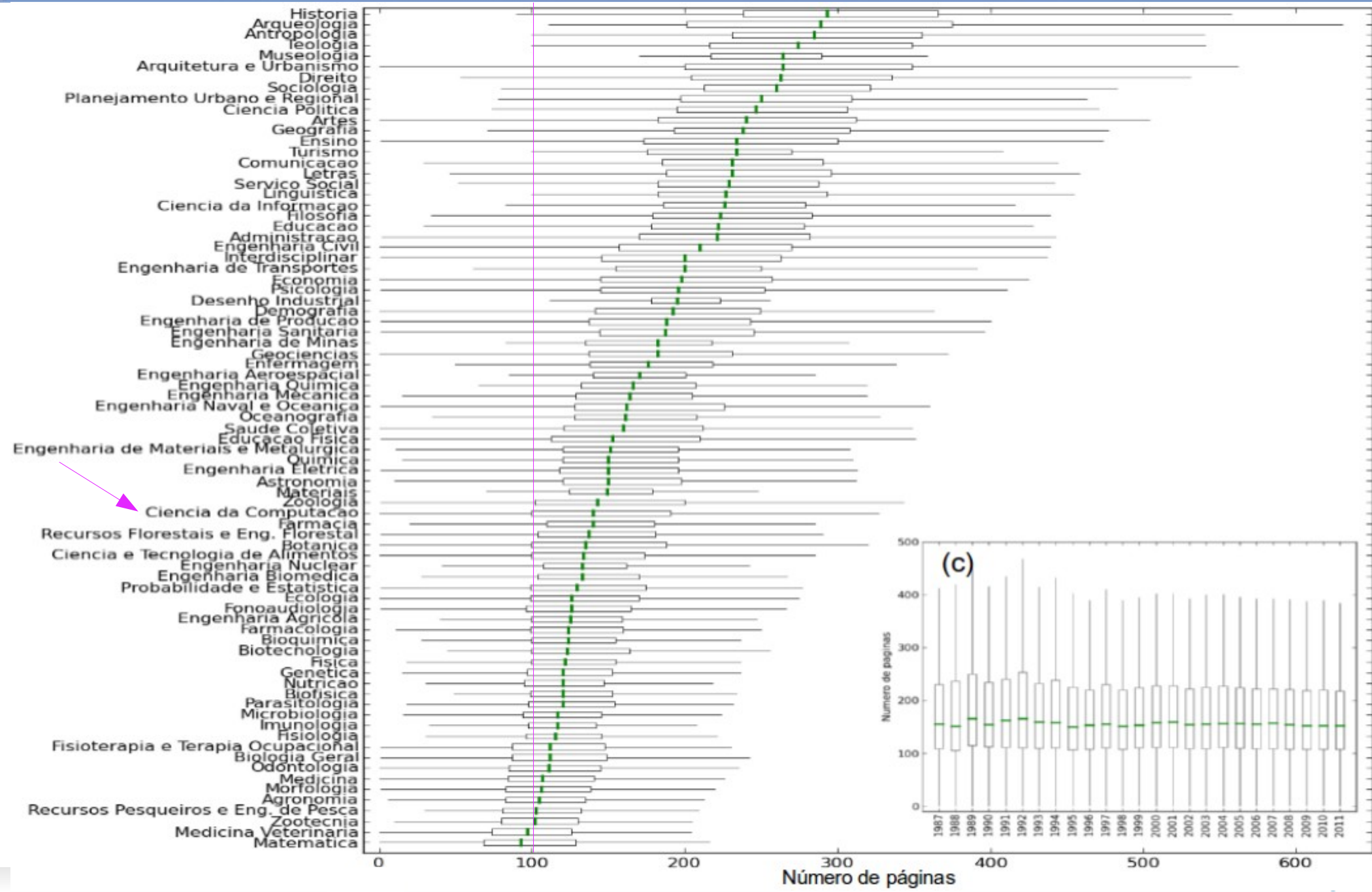


**Quantas páginas devo de escrever para  
minha monografia?**

# Dissertação



# Tese







# **Sobre os acompanhamentos de defesa**

# Templates

## 1 Identificação do documento

- Título do documento: Texto texto texto [1]
- Nome completo do autor principal: Texto texto texto
- Área de pesquisa do autor principal: Texto texto texto
- Número de citações do documento: Número
- Índice-h do autor principal: Número

## 2 Liste duas perguntas que o documento tenta responder

- Texto texto texto texto ... texto ?
- Texto texto texto texto ... texto ?

## 3 Liste duas publicações ‘importantes’ que citaram o documento

- Título da publicação A [2]
- Título da publicação B [3]

## 4 Resumo (limite aproximado de 500 palavras)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Fusce nec urna tempus, maximus libero vel, lacinia tellus. Mauris ut dolor odio. Aliquam diam ligula, pellentesque eu pretium eu, eleifend ut lectus. Suspendisse placerat suscipit luctus. Duis nec eros mattis, pulvinar felis sed, ultricies odio. Quisque sit amet mi at augue tempor tincidunt quis at enim. In molestie nec magna vitae efficitur. Aliquam nibh felis, fermentum eu pharetra at, ornare id leo. Nullam mi nulla, facilisis at enim id, auctor pretium lacus. Cum sociis natoque penatibus et magnis dis parturient montes, nascetur ridiculus mus. Fusce venenatis suscipit dolor nec mollis. Ut sed diam libero.

Vestibulum interdum leo ac pellentesque lobortis. Maecenas dictum, magna ac laoreet iaculis, ipsum dolor consequat dolor, sed convallis augue lacus vel mauris. Nunc sed orci eu nunc volutpat placerat. Praesent auctor cursus eros sit amet sollicitudin. Donec egestas diam id augue tristique placerat. Etiam bibendum nisi lacus, vel finibus ipsum viverra ac. Nulla odio quam, mattis nec maximus sit amet, tristique a lorem. Praesent sagittis, diam vitae posuere vehicula, purus tellus eleifend quam, vitae viverra neque diam nec sem. Sed pulvinar enim viverra, tincidunt mi at, sodales erat. Etiam consequat,

## Referências

- [1] Uri Alon. How To Choose a Good Scientific Problem. *Molecular Cell*, 35(6):726–728, September 2009.
- [2] Umberto Eco. *Como se Faz uma Tese*. Perspectiva, 22ª edition, 2009. Tradução Gilson Cesar Cardoso de Souza.
- [3] Justin Zobel. *Writing for Computer Science: The art of effective communication*. Springer, segunda edition, 2004.

## 1 Identificação da defesa

- Título do trabalho: Texto texto texto ... texto
- Nível de defesa: Mestrado/doutorado
- Nome do apresentador(a): Texto texto texto ... texto
- Nome do(a) orientador(a): Texto texto texto ... texto
- Instituição: Texto texto texto ... texto
- Data e local de defesa: Texto texto texto ... texto
- Número de pessoas na sala: Número aproximado

## 2 Sobre o tempo de defesa (em minutos)

- Apresentação do aluno: Número aproximado
- Participação do avaliador 1: Número aproximado
- Participação do avaliador 2: Número aproximado
- ...
- Participação do avaliador N: Número aproximado
- Tempo total da defesa: Número aproximado

## 3 Perguntas importantes

### 3.1 Duas perguntas ‘fáceis’ realizadas pelos avaliadores

- Texto texto texto texto ... texto ?
- Texto texto texto texto ... texto ?

### 3.2 Duas perguntas ‘difíceis’ realizadas pelos avaliadores

- Texto texto texto texto ... texto ?
- Texto texto texto texto ... texto ?

### 3.3 Duas perguntas que, segundo seu parecer, deveriam ser realizadas na defesa

- Texto texto texto texto ... texto ?
- Texto texto texto texto ... texto ?

### 3.4 Qual é a maior contribuição do trabalho apresentado? (opcional)

Lorem ipsum dolor sit amet, consectetur adipiscing elit. Fusce nec urna tempus, maximus libero vel, lacinia tellus. Mauris ut dolor odio. Aliquam diam ligula, pellentesque eu pretium eu, eleifend ut lectus.