



Escrita e Apresentação de Artigos Científicos

Mirella M. Moro - mirella@dcc.ufmg.br
EVCOMP 2018 @ PPGCC, DCC, UFMG

www.dcc.ufmg.br/~mirella



MIRELLA M. MORO

[HOME](#) / [CALL FOR PAPERS](#) / [ALUNOS](#)

CONTACT INFORMATION

Mirella M. Moro

Departamento de Ciência da Computação
Universidade Federal de Minas Gerais - UFMG
Av. Antônio Carlos, 6627
Pampulha, Belo Horizonte - MG, Brazil
CEP 31270-010
Office: Building ICEX/DCC, room 6326
Phone numbers:
+55 (31) 3409-5849 (Office)
+55 (31) 3409-5858 (Fax)
Email: mirella [AT] dcc [DOT] ufmg [DOT] br



Introdução

...2017

Antes da Escrita
Escrita: Início
Escrita: Centro
Escrita: Fim
Finalmente

2018

Processo de Escrita
Estrutura
Linguagem
Estilo
Apresentação
Finalmente

ROTEIRO

1. Processo de Escrita
2. Estrutura
3. Linguagem
4. Estilo
5. Apresentação
6. Finalmente



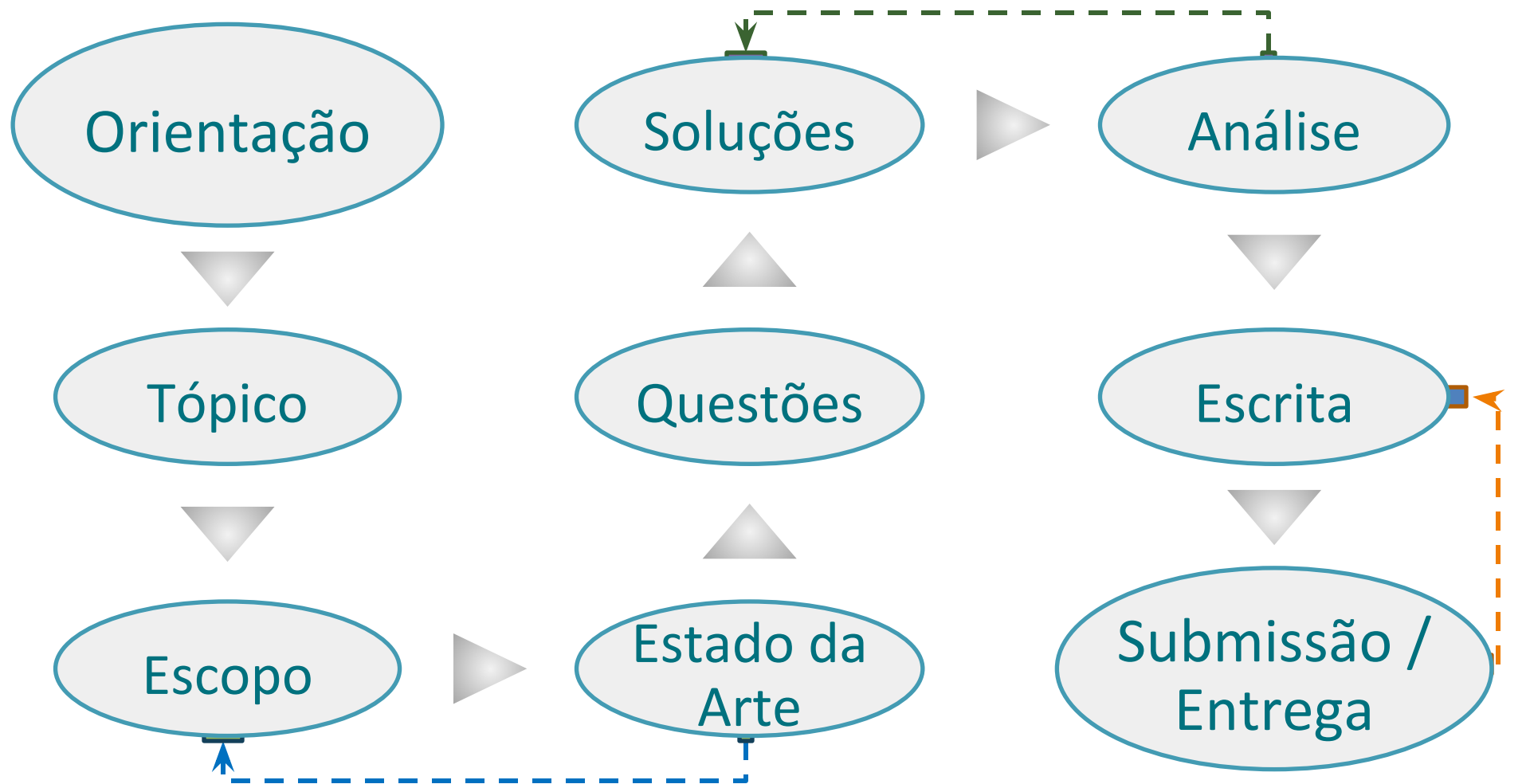
Dica

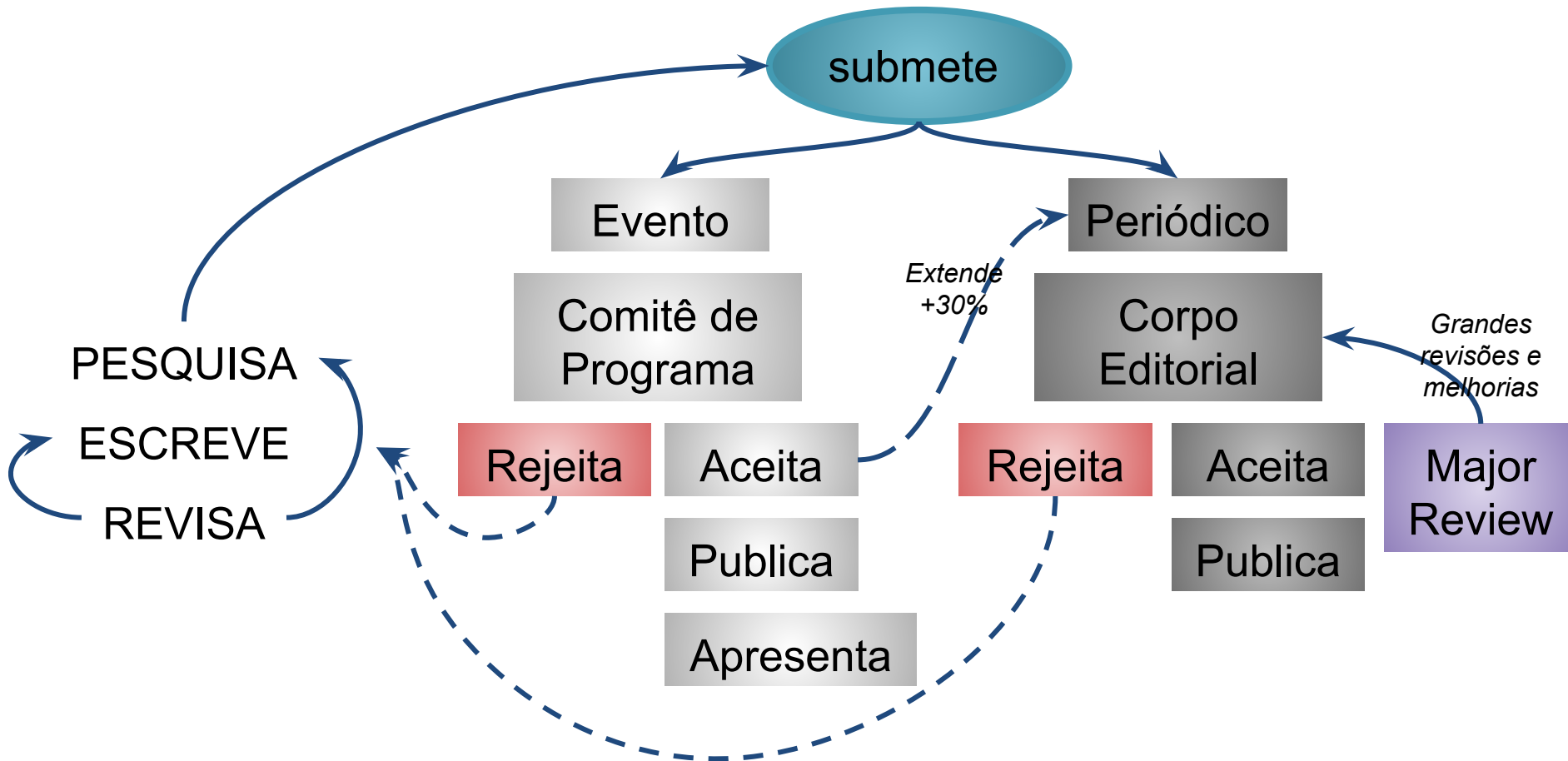
Artigos Científicos
Conferências
Periódicos
Monografia
Dissertação
Tese
Relatório Técnico
Trabalho p/ Disciplina

Processo de Escrita

PARTE 1









**Pesquisa
+ Escrita**



Planejamento



As Complexidades da Tarefa

- Como tornar um **interesse vago** em um problema **merecedor de apresentação** e solução
- Como construir um **argumento** que motiva leitores a **aceitar** o que você diz
- Como antecipar **dúvidas** de atentos mas **críticos** leitores e como respondê-las apropriadamente
- Como criar uma **introdução e conclusão** que respondam à pergunta mais difícil delas, **E eu com isso?**
- Como **ler** sua própria escrita como outros irão, e então aprender quando e como **revisá-la**

Receita de Planejamento

Defina o OBJETIVO

Tópico: Estou pesquisando _
Questão: a fim de descobrir o
que/como/por que _
Relevância: para _
Aplicação prática potencial:
que pode ser utilizada em _

Defina o ESPECTRO

< 4 palavras:
muito abrangente

Entenda o PÚBLICO ALVO

Quem?
Background?
Expectativas?

Objetivo+Espectro+Público: Exemplo Real



ACM SAC/ CSBC	• XML Database • XML Query Processing • RoXSum
SIGMOD/ SBBD	• XML Query Filtering • XML Routing • RoXSum
WebDB	• XML Routing • RoXSum • Algorithms

**+GENÉRICO
+BACKGROUND**



**+ESPECÍFICO
+DETALHES**

Planejamento

Onde? O que? Por quê? Para quê? Como?

Quanto? “largura e fundura”

Quem irá ler seu texto?

Eles esperam que faça o que pretende fazer? contexto, novidade, contribuição, surpresa

Quanto eles já sabem?

Como responderão aos problemas e soluções apresentados?

Enquanto estiver

Escrevendo
Apresentando
Pensando
Planejando
Discutindo
Reunindo
Argumentando

con

CONTEXTO



pro

PROBLEMA



s

SOLUÇÃO



a

AVALIAÇÃO



contexto

Contexto geral

Contexto específico [estado da arte]



problema

O que *não* existe/funciona?

O que pode melhorar?



solução

Contribuição

Como resolver o problema



avaliação

É bom? É melhor?

[Nem] Sempre: Quando? Por quê?



Contribuições por tudo (!)

contexto

$X \text{ atual} > \text{Novo contexto}$



problema

$X \text{ atual} > \text{Novo problema}$



solução

$X \text{ atual} > \text{Nova solução (+comum)}$



avaliação

$\text{Diferentes avaliações} > X \text{ atual}$



Onde $X \in \{ \text{contexto, problema, solução, avaliação} \}$

contexto

Dados relacionais > streams
Dados relacionais > móveis



problema

Dados relacionais > semiestruturado
Modelo relacional > big data



solução

+ COMUM



avaliação

Escolha qualquer um e avalie em: paralelo,
distribuído, grandes volumes, map-reduce, novos hardwares, ...



Exemplo [ZHANG et al @ SIGMOD 1996]

Contexto	<u>Finding useful patterns</u> in large datasets has attracted considerable interest recently,
Problema	and <u>one of the most widely studied problems in this area is</u> the identification of clusters, or densely populated regions, in a multi-dimensional dataset. <u>Prior work does not</u> adequately address the problem of large datasets and minimization of I/O costs. <u>This paper presents</u> a data clustering method named BIRCH (Balanced Iterative Reducing and
Solução	Clustering using Hierarchies), and demonstrates that it is especially suitable for very large databases. BIRCH incrementally and dynamically clusters incoming multi-dimensional metric data points to try to produce the best quality clustering with the available resources (i.e., available memory and time constraints). BIRCH can
(Detalhes)	typically find a good clustering with a single scan of the data, and improve the quality further with a few additional scans. BIRCH is also the first clustering algorithm proposed in the database area to handle “noise” (data points that are not part of the underlying pattern) effectively. <u>We evaluate</u> BIRCH’s time/space efficiency, data input order sensitivity, and clustering quality through several experiments. We also present a
Avaliação	<u>performance comparison</u> of BIRCH versus CLARANS, a clustering method proposed recently for large datasets, and show that BIRCH is consistently superior.

Não é só em BD: RAGHAVAN et al @ SIGCOMM 2007

Contexto	Today's cloud-based services integrate globally distributed resources into seamless computing platforms. Provisioning and accounting for the resource usage of these Internet-scale applications presents a challenging technical problem. This paper presents the design and implementation of distributed rate limiters, which work together to enforce a global rate limit across traffic aggregates at multiple sites, enabling the coordinated policing of a cloud-based service's network traffic. Our abstraction not only enforces a global limit, but also ensures that congestion-responsive transport-layer flows behave as if they traversed a single, shared limiter. We present two designs—one general purpose, and one optimized for TCP—that allow service operators to explicitly trade off between communication costs and system accuracy, efficiency, and scalability. Both designs are capable of rate limiting thousands of flows with negligible overhead (less than 3% in the tested configuration). We demonstrate that our TCP-centric design is scalable to hundreds of nodes while robust to both loss and communication delay, making it practical for deployment in nationwide service providers.
Problema	
Solução	
(Detalhes)	
Avaliação	

Escrita Científica

PRONTO

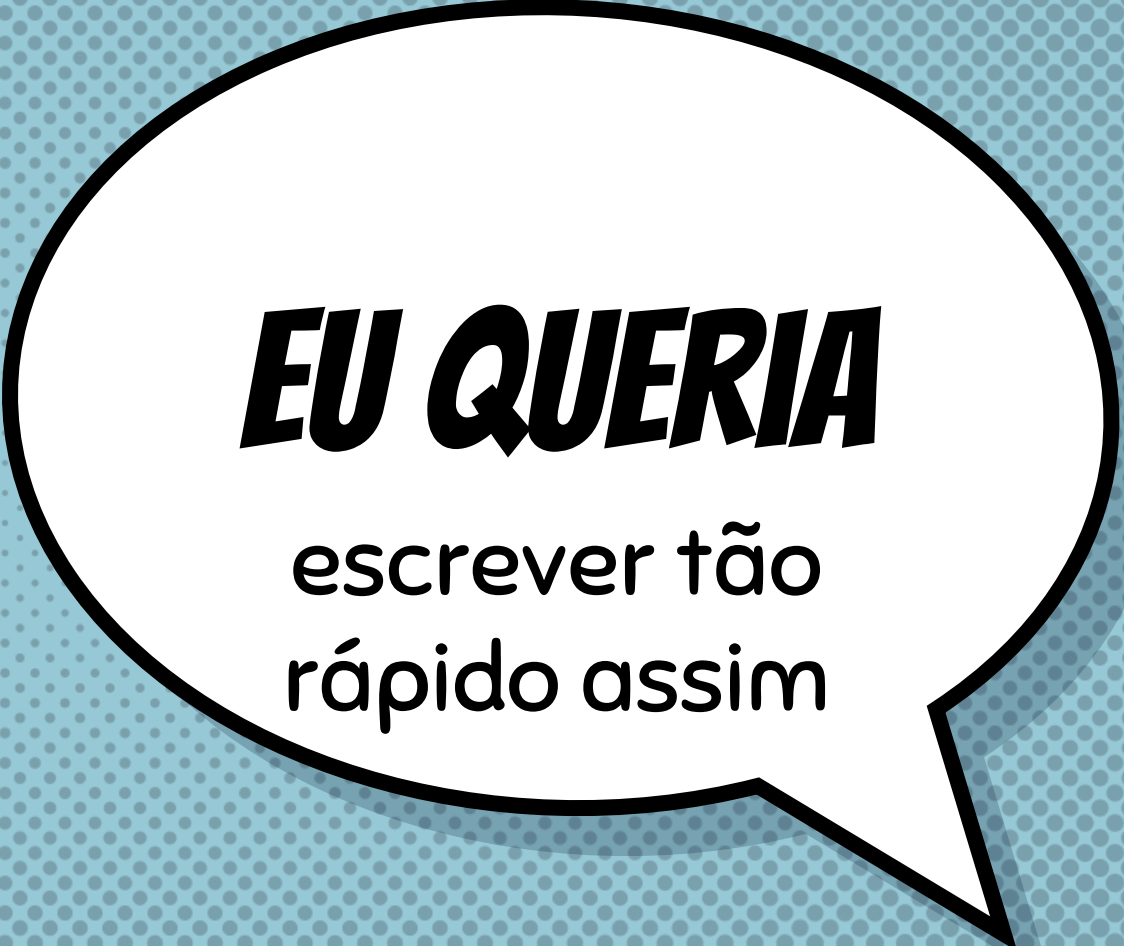
- Título
- Autor (es)
- Resumo
- Introdução
- Corpo
- Conclusão
- Referências

A FAZER

- Título
- Autor (es)
- Resumo
- Introdução
- Corpo
- Conclusão
- Referências



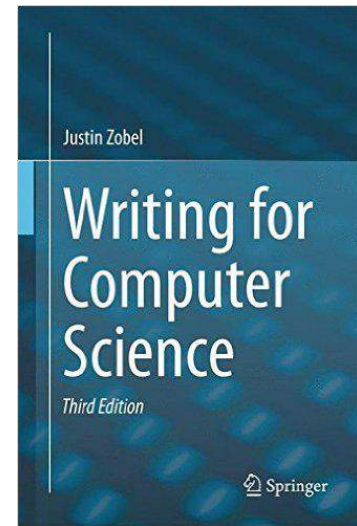
ergo
este curso



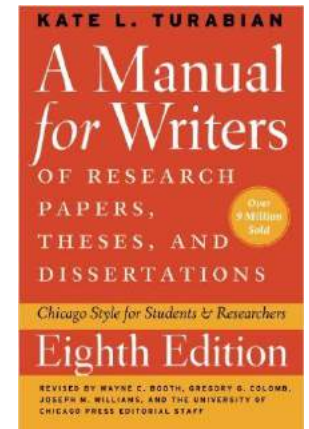
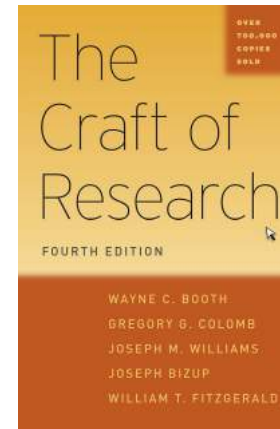
EU QUERIA

escrever tão
rápido assim

**Practice
Makes
Perfect**



~US\$40



~US\$18

ROTEIRO

1. Processo de Escrita ✓ ✓
2. Estrutura
3. Linguagem
4. Estilo
5. Apresentação
6. Finalmente



Dica

Artigos Científicos
Conferências
Periódicos
Monografia
Dissertação
Tese
Relatório Técnico
Trabalho p/ Disciplina

Estrutura

PARTE 2



Estrutura Base

ARANTES @ SBBD 2003

1. Introdução
2. Trabalhos Relacionados
3. Motivação e Conceitos Fundamentais
4. Composição de Operadores por Similaridade: os Novos Algoritmos
5. Experimentos Realizados
6. Conclusões e Trabalhos Futuros

BRAGANHOLO @ CTD 2005

1. Introduction
2. Related Work
3. Query Trees
4. Update Language
5. Mapping
6. Summary and Concluding Remarks

Estrutura Base

M. TURFANO et al @ ICSE 2015

1. Introduction
2. Study Design
3. Analysis of the Results
4. Threats to Validity
5. Related Work
6. Conclusion and Lessons Learned

L. LIN et al @ SIGCOMM 2004

1. Introduction
2. Background and Related Work
3. A First Principles Approach
4. Topology Metrics
5. Comparing Topologies
6. Discussion

Estrutura Base

V. IYER et al @ SIGCOMM 2016

1. Introduction
2. System Design
3. FPGA and IC Design
4. Evaluation
5. Proof-of-Concept Applications
6. Related Work
7. Discussion and Conclusions

VASISHT et al @ SIGCOMM 2016

1. Introduction
2. Related Work
3. Background
4. Intuition Underlying R2-F2
5. Algorithm
6. Integrating R2-F2 with the LTE Architecture
7. Implementation
8. Results
9. Concluding Remarks

Estrutura Base

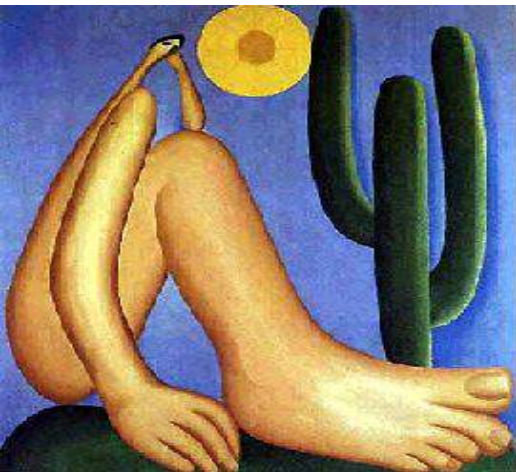
CHANG et al@ ACM TOCS 08

1. Introduction
2. Data Model
3. API
4. Building Blocks
5. Implementation
6. Refinements
7. Performance Evaluation
8. Real Applications
9. Lessons
10. Related Work
11. Conclusions
12. Acknowledgements
13. References

DEAN & GHEMAWAT @ CACM 08

1. Introduction
2. Programming Model
3. Implementation
4. Refinements
5. Performance
6. Experience
7. Related Work
8. Conclusions
9. Acknowledgements
10. References

Escrever Artigos como Arte



Estrutura Base

Mesmo com exceções, funciona em
~90% dos trabalhos de Computação

Título

Autor (es)

Resumo

Introdução

♥Corpo♥

Conclusão

Referências

Título



<https://pixabay.com>

- Referência principal ao trabalho
- Chave para ser referenciado
- Claro, curto, correto
 - Nome, não uma frase, original
 - Primeira coisa a se escrever??

Exemplos

- **Hyperspectral-based predictive modelling of grapevine water status in the Portuguese Douro wine region.** Pôças et al: IJAEOG, 2017.
- **What's on the grapevine?** Angel et al: SIGMOD, 2009.
- **"I Heard It on the Grapevine" - Blogging, Facebook, YouTube, and Student Self-organization during a Faculty Strike.** Gould: HCI, 2009.
- **Maximum Likelihood from Incomplete Data via the EM Algorithm.** Dempster, Laird & Rubin, JSTOR 1977. [52k+ cites](#)
- **Distinctive image features from scale-invariant keypoints.** Lowe, Int'l J. Comp.Vision, 2004. [45k+ cites](#)
- **A Method for Obtaining Digital Signatures and Public- Key Cryptosystems.** Rivest, Shamir & Adleman, CACM 1978. [18k+ cites](#)
- **YACC:** Yet another compiler-compiler. S.C. Johnson, 1979
- **Wander Join:** Online aggregation via Random Walks. Li et al @ SIGMOD 2016

Autor (es)

- Nome completo (“artístico”) + filiação + email
- Ordem dos autores??
 - Não há regra padrão aceita globalmente

- **André Silva, Bento Muniz, Carla Costa**
- **Carla Costa, Bento Muniz, André Silva**
- **Bento Muniz, Carla Costa, André Silva**

Aluno “dono” tese

Bolsista

Orientador

Resumo

Um parágrafo 100-250 palavras (~ # páginas)

– Propaganda ou trailer do artigo

– Atrai (?) a atenção e o interesse do leitor

- Sempre menciona informações ou conclusões do texto
- Sem referências bibliográficas (exceto em ocasiões raras, como modificações a um método publicado previamente)
- Primeiro a ser escrito: planejamento
- Último a ser escrito: ideia melhor

Resumo

Contexto

Problema

Solução

Avaliação



Resumo em estilo curto

1 OU 2 LINHAS PARA CADA ITEM

Contexto	<u>Structural summaries</u> are data structures that preserve all structural features of XML documents in a compact form.
Problema/ Solução	<u>We investigate the applicability</u> of the most popular summaries as access methods within XML query processing. In this context, <u>issues</u> like space and false positives introduced by the summaries need to be
Avaliação	<u>examined</u> . <u>Our evaluation</u> reveals that the additional space required by the more precise structures is usually small and justified by the considerable performance gains that they achieve.

MORO et al – WWW 2006

Contexto Geral

Contexto Específico

Problema

Est-da-arte

Solução

Detalhes

Avaliação

Publish-subscribe applications are an important class of content-based dissemination systems where the message transmission is defined by the message content, rather than its destination IP address. **With the increasing use of XML as the standard format on many Internet-based applications, XML aware pub-sub applications become necessary. In such systems, the messages (generated by publishers) are encoded as XML documents, and the profiles (defined by subscribers) as XML query statements. As the number of documents and query requests grow, the performance and scalability of the matching phase (i.e. matching of queries to incoming documents) become vital. Current solutions have limited or no flexibility to prune out queries in advance.** In this paper, we overcome such limitation by proposing a novel early pruning approach called Bounding-based XML Filtering or BoXFilter. **The BoXFilter is based on a new tree-like indexing structure that organizes the queries based on their similarity and provides lower and upper bound estimations needed to prune queries not related to the incoming documents.** Our experimental evaluation shows that the early profile pruning approach offers drastic performance improvements over the current state-of-the-art in XML filtering.

MORO et al @ VLDB 2007

Contexto

Middleboxes are ubiquitous in today's networks and perform a variety of important functions, including IDS, VPN, firewalling, and WAN optimization. These functions differ vastly in their requirements for hardware resources (e.g., CPU cycles and memory bandwidth). Thus, depending on the functions

Problema

Estado da arte

they go through, different flows can consume different amounts of a middlebox's resources. While there is much literature on weighted fair sharing of link bandwidth to isolate flows, it is unclear how to schedule *multiple* resources in a middlebox to achieve similar guarantees. In

Contribuição

this paper, we analyze several natural packet scheduling algorithms for multiple resources and show that they have undesirable properties. We propose a new algorithm, Dominant Resource Fair Queuing (DRFQ), that retains the attractive properties that fair sharing provides for one resource.

Relevância

In doing so, we generalize the concept of virtual time in classical fair queuing to multi-resource settings. The resulting algorithm is also applicable in other contexts where several resources need to be multiplexed in the time domain.

GHODSI et al – SIGCOMM 2012

Contexto é um tópico

Relevância

Contribuição (topologia!)

We have had to wait over 30 years since the naive Bayes model was first introduced in 1960 for the so-called Bayesian network classifiers to resurge. Based on Bayesian networks, these classifiers have many strengths, like model interpretability, accommodation to complex data and classification problem settings, existence of efficient algorithms for learning and classification tasks, and successful applicability in real-world problems. In this article, we survey the whole set of discrete Bayesian network classifiers devised to date, organized in increasing order of structure complexity: naive Bayes, selective naive Bayes, seminaive Bayes, one-dependence Bayesian classifiers, k -dependence Bayesian classifiers, Bayesian network-augmented naive Bayes, Markov blanket-based Bayesian classifier, unrestricted Bayesian classifiers, and Bayesian multinets. Issues of feature subset selection and generative and discriminative structure and parameter learning are also covered.

BIELZA & LARRAÑAGA – COMP. SURVEYS 2014

mirella@dcc.ufmg.br

Introdução

**UM ARTIGO CIENTÍFICO NÃO É
UM *LIVRO DE SUSPENSE* NO
QUAL O LEITOR SÓ DESCOBRE
O QUE ESTÁ ACONTECENDO
NO FINAL, *OU NÃO***

—

Erro mais comum

Deixar tudo implícito

Achar que é óbvio



Dica

O óbvio
não é ciência,
é senso comum

Introdução

- Uma introdução bem escrita é fundamental!!
- O leitor deve estar ciente do que acontece desde o início, **desde a introdução** *(na verdade, desde o Resumo)*
- A introdução é uma reafirmação estendida do conteúdo do Resumo (linhas→parágrafos)

Introdução

Contexto

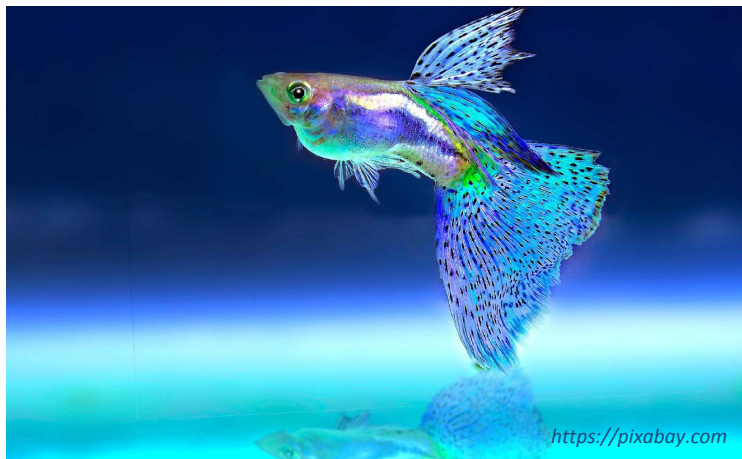
Problema

Solução

Avaliação



Introdução contém em algum parágrafo



- **Como** o seu trabalho pode ser empregado
- **Onde** o seu trabalho pode ser empregado, quais os contextos, **quais** aplicações podem se beneficiar
- Qual a **contribuição social**
- Quais **problemas práticos** o seu trabalho resolve

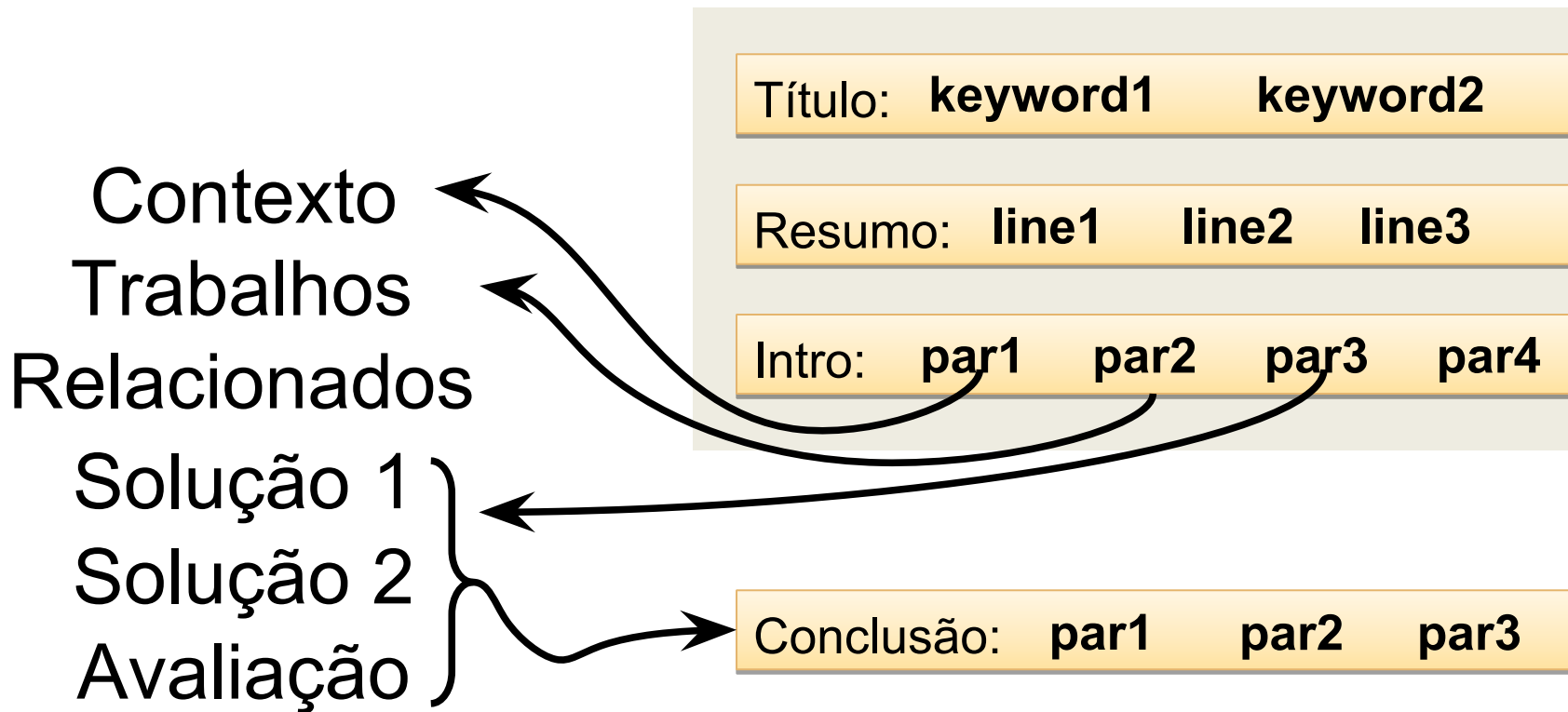
Yesterday's version of distributed computing was a selfcontained, colocated server farm. **Today**, applications are increasingly deployed on third-party resources hosted across the Internet. Indeed, the rapid spread of open protocols and standards like Web 2.0 has fueled an explosion of compound services that script together third-party components to deliver a sophisticated service [27, 29]. These specialized services are just the **beginning**: flagship consumer and enterprise applications are increasingly being delivered in the software-as-a-service model [9]. **For example**, Google Documents, Groove Office, and Windows Live are early examples of desktop applications provided in a hosted environment, and represent **the beginning of a much larger trend**.

One of the **key barriers** to moving traditional applications to the cloud, however, is the loss of cost control [17]. In the cloud-based services model, cost recovery is typically accomplished through metered pricing. Indeed, **Amazon's EC2** charges incrementally per gigabyte of traffic consumed [3] [...] Limiting global resource consumption in a distributed environment, however, presents a **significant technical challenge**. **Ideally**, resource providers would not require services to specify the resource demands of each distributed component *a priori*; *such fine-grained measurement and modeling can be challenging* for rapidly evolving services. **Instead**, they should provide a fixed price for an aggregate, global usage, and allow services to consume resources dynamically across various locations, subject to the specified aggregate limit. [...]

This paper makes three primary contributions:

- **Rate Limiting Cloud-based Services**. We identify a key challenge...
- **Distributed Rate Limiter Design**. We present the design and ...
- **Evaluation and Methodology**. We develop a methodology...

Até aqui ...



Corpo Central

1. Introdução
2. JÁ EXISTE
3. NOVIDADE
4. VALIDAÇÃO
5. Conclusão

>>> Sempre olhe o estilo de artigo publicado nos principais veículos da sua área específica

O que já Existe

> Já existe

Novidade
Validação
Discussão



<https://pixabay.com>

Conceitos Básicos +
Trabalhos Relacionados

Apresentados juntos/não

Conceitos antes da contribuição

Relacionados no início ou fim

Já existe: Conceitos Básicos

O que é
necessário para
entender o
trabalho e suas
contribuições?

**Definições
Notações**

Modelos

**Arquitetu
ras**

Linguagens

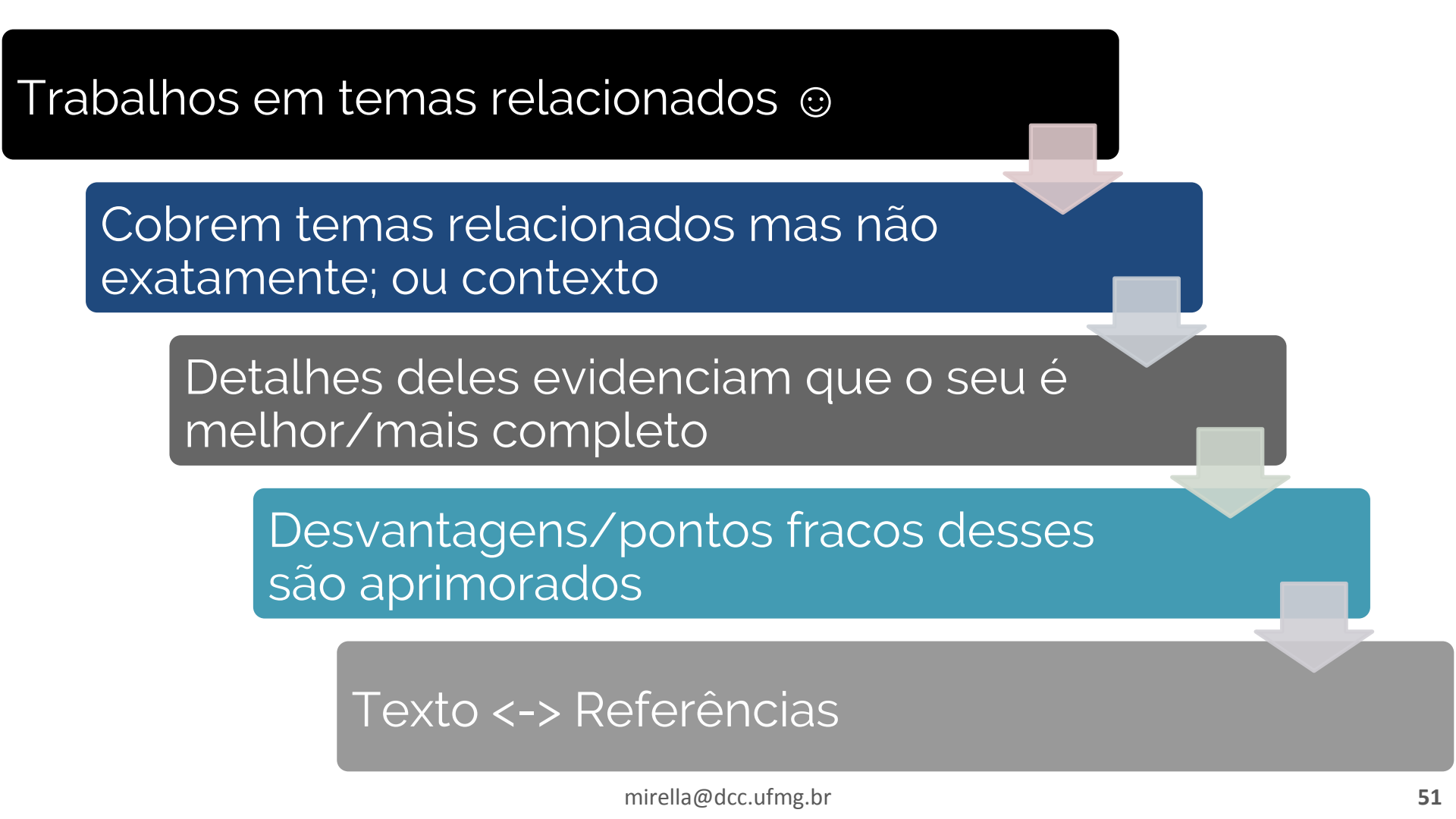
Cenários

Padrões

Já Existe: Trabalhos Relacionados



Trabalhos em temas relacionados ☺



Cobrem temas relacionados mas não exatamente; ou contexto

Detalhes deles evidenciam que o seu é melhor/mais completo

Desvantagens/pontos fracos desses são aprimorados

Texto <-> Referências

Já Existe: Trabalhos Relacionados

- **Seção de Trabalhos Relacionados é praticamente obrigatória**
- **Como o seu trabalho se relaciona a todos os mencionados**
- Por exemplo, o título é:

“Resumos Estruturais em Processamento de Consulta XML”

Exemplos de temas relacionados:

- Resumos estruturais em outros contextos (outros processamentos)
- Resumos estruturais para processar outros tipos de dados
- Outras formas de processar consultas XML
- ...



Novidade

Contribuições

Já existe
> **Novidade**
Validação
Discussão

- ★ Parte principal do artigo
- ★ A que veio?
- ★ O que adiciona?
- ★ Tem de estar claríssimo

Novidade

- Parágrafo com ideia geral da proposta
- Esclareça novas definições
 - Claramente: são novas definições do artigo
- Adicione quantos parágrafos necessários para apresentar:
 - o que é o trabalho
 - como funciona a proposta
 - o que é novidade, por que
 - detalhes e explicações sobre partes principais do funcionamento da proposta



BIG

NOVIDADE



O que

Por que

Como

Onde

Quanto

+ Impactos (sociais, econômicos, etc)

Validação, Avaliação

Já existe
Novidade
> **Validação**
Discussão

Solução proposta funciona e seus benefícios

- Análise
- Estudo de caso
- Experimentos



Validação: **Análise**

Proposta é correta (demonstração e provas)

Inclua (conforme necessário):

- Um parágrafo com o resumo do que é provado nessa seção
- Um parágrafo com definições específicas usadas na análise (ex. estruturas usadas nas provas)
- Provas e análises
- Comentários finais sobre o significado das provas de um modo intuitivo ou num nível mais prático

Validação: **Estudo de Caso**

Proposta é aplicável, implementável

Inclua (conforme necessário):

- Contexto geral
- Regras ou condições necessárias nesse estudo de caso
- Modelagem / Implementação
- Funcionamento
- Vantagens e desvantagens de usar o modelo proposto nesse estudo de caso

Validação: **Experimentos**

Proposta funciona, desempenho bom/superior

Inclua (conforme necessário):

- Contexto: o que é medido, o que não é, por que
- Modelo de simulação ou infraestrutura das medições: configuração do sistema, tipo de máquinas usadas, linguagens, ...
- Resultados dos experimentos [próximo slide]
- Comentários finais, discussões, explicações adicionais

Validação: Experimentos

Descrição de resultados

- **Subtítulo** (em negrito): para diferenciar experimentos (ex: avaliação do tamanho da entrada, variando a quantidade de consultas, dados sintéticos versus dados reais, ...)
- **Razão clara**: por que cada gráfico aparece no seu artigo (ex: conforme o tamanho dos arquivos de entrada aumenta, o throughput diminui, ...)
- **Explicar**: os eixos dos gráficos, o que o gráfico mostra, qual é a tendência, por que a tendência aparece, por que um algoritmo apresenta melhores resultados que outro, ...
- **Auto-contido**: legendas devem ser legíveis e compreensíveis e suficientes para entender o gráfico

Experimentos: Exemplo

6. Performance studies

6.1 Analysis

6.2 Synthetic Dataset Generator

6.3 Parameters and Default Setting

6.4 Base Workload Performance

6.5 Sensitivity to Parameters

6.6 Time Scalability

6.7 Comparison of BIRCH and CLARANS

6.8 Application to Real Datasets **ZHANG @ SIGMOD 1996**



Discussão

Já existe
Novidade
Validação
> **Discussão**

- Como subseção final de Experimentos ou na Conclusão
- Relacionamentos entre os fatos e resultados observados
- Princípios, relações, generalizações mostrados nos Experimentos
- Exceção ou falta de relação, pontos incertos
- Resultados e interpretações concordam (ou contrastam) com relacionados
- Implicações teóricas e possíveis aplicações práticas

CONCLUSÃO

Sugestão: ser mais específico que na introdução e informar (um parágrafo/linha por item)

- resumo do que o artigo apresentou
- principais resultados e contribuições, com devidas evidências
- comentários sobre a importância, relevância ou
- dicas para o uso prático do seu trabalho (como os resultados dos experimentos podem ajudar na prática...)
- trabalhos futuros

Evidências para cada conclusão (não assuma que o leitor é super capaz de juntar todos os pontos sozinho)

Conclusão: exemplo

“As cloud-based services transition from marketing vaporware to real, deployed systems, the demands on traditional Web-hosting and Internet service providers are likely to shift dramatically. **In particular**, current models of resource provisioning and accounting lack the flexibility to effectively support the dynamic composition and rapidly shifting load enabled by the software as a service paradigm. **We have identified one key** aspect of this problem, namely the need to rate limit network traffic in a distributed fashion, and provided two novel algorithms to address this pressing need.

Our experiments show that naive implementations based on packet arrival information are unable to deliver adequate levels [...]

Our results demonstrate that it is possible to recreate, at distributed points in the network, the flow behavior that end users and network operators expect from a single centralized rate limiter. Moreover, it is possible [...]” **RAGHAVAN et al - SIGCOMM 2007**

REFERÊNCIAS

- Corretas, completas, específicas
- Informações obrigatórias: autores, título, nome do evento ou periódico (editora), volume e número se necessário, ano
- Referências relevantes
 - Do mesmo ano (ou ano anterior) para ilustrar que o tópico é atual e de interesse da comunidade
 - Artigos de conferências, periódicos, livros (não apenas sites da Internet!)
 - Todas as obras listadas no conjunto de referências devem ser mencionadas no texto, e vice-versa

Até Agora

Estrutura

{Título - Autores - Resumo}

Introdução

O que já existe

Novidade

Validação

Conclusão

{Referências}

Contexto

Problema

Solução

Avaliação

ROTEIRO

1. Processo de Escrita ✓ ✓
2. Estrutura ✓ ✓
3. Linguagem
4. Estilo
5. Apresentação
6. Finalmente



Dica

Artigos Científicos
Conferências
Periódicos
Monografia
Dissertação
Tese
Relatório Técnico
Trabalho p/ Disciplina

Linguagem

PARTE 3



Sete Pecados Capitais

Português, inglês,
qualquer texto
(mesmo não
científico)





#1 Frase Longas

Repletas de vírgulas ou não

É pior uma frase longa sem vírgula ou a mesma com dez vírgulas?

Se tiver de ler três vezes para entender, já era

Regra simples: no máximo duas vírgulas

Se precisar mais: considere dividir a sentença



#2 Erros Ortográficos

Um erro ortográfico pode distrair o leitor

10 ou mais: leitor irá se perder a ponto de ignorar o seu trabalho ou **rejeitá-lo** (se for um avaliador)



#3 Erros Gramaticais

Paralelismo, concordância e conjugação

Língua portuguesa: crase



#4 Tradução Literal

Traduzir cada palavra literalmente para o inglês não funciona

tea with me não faz sentido

Imbromation (inventar palavras em inglês). Na dúvida, consulte um bom **dicionário** para verificar se a palavra existe, ou um thesaurus para verificar se um sinônimo fica mais adequado



#5 Elementos Ilegíveis

Se a imagem, gráfico, esquema, modelagem, tabela, etc. precisa aparecer no texto, **então precisa estar legível**

Na dúvida, imprima o texto no seu tamanho final e verifique



#6 Cópia Literal

Copyright?

Ética?

Plágio?



#7 Blablabla

Encher linguiça

Se o limite do artigo é 12 páginas mas você só consegue escrever 10, não vá sair escrevendo qualquer bobagem e chovendo no molhado para chegar às 12

O avaliador saberá que você está enchendo linguiça, pode acreditar

Peculiaridades da Escrita Científica

Não é redação do
ENEM...





Tempo Verbal

Sempre no presente

- ✗ In this paper, we will explain how ABC...
- ✓ Next, we explain how ABC works...
- ✗ The next sections will describe the...
- ✓ Section 3 describes the algorithm...
- ✗ First, we have run the algorithm in a Pentium... and chose to evaluate it through a series of experiments... [cenários].
- ✓ Next, we describe the evaluation procedure for comparing our algorithm against the baselines. The evaluation setup considers a Pentium ... Then, the series of evaluations contains: ...[cenários].



Tempo Verbal

Sempre no presente

Quatro exceções

1. (pode) passado para trabalhos anteriores: **Moro mostrou que ... [20].**
2. (pode) passado nas seções finais de cada capítulo: **Este capítulo apresentou as definições sobre...**
3. (pode) passado nas conclusões: **Este trabalho introduziu uma ...**
4. futuro apenas para trabalhos que serão realizados [na Conclusão]: **Como trabalhos futuros, serão definidas novas formas de...**



Primeira Pessoa

Do singular é muito raro

Trabalhos em Computação geralmente são realizados pelo estudante e seu orientador, ou professor de disciplina, ou por times de P&D $\Rightarrow$ raro ter trabalhos científicos de pesquisa na primeira pessoa do singular

Mesmo monografia, dissertação e tese: com o orientador

Se individualmente:

- ✘ Neste trabalho, eu apresento um novo algoritmo para ...
- ☑ Neste trabalho, apresenta-se um novo algoritmo para ...
- ☑ Neste trabalho, é apresentado um novo algoritmo para ...



Primeira Pessoa

Do plural refere-se aos **autores**

muito muito muito **CUIDADO** !!!

- We make summation of the number of ...
- Our approach is ABC, in which we will use fewer bits for characters...
- The experimental evaluation we are seeking emphasize time, space and memory consumption.
- ... the properties that we use in our analysis (Section 5).
- Formally, we have that $PSC(A,B) = \dots$
- If we have a limited alphabet, ...

Corrige-se para

- + The approach sums the number of...
- + Our approach is ABC that uses fewer bits for characters...
- + The experimental evaluation emphasizes time, space and memory consumption.
- + ... the properties used in Section 5.
- + Formally, $PSC(A,B)$ is given by ...
- + If there is a limited alphabet, ...



Voz Passiva

Foco no objeto que sofreu a ação

Deve ser **evitada** ao máximo



Local

Em cima, embaixo, puxa e vai

- A figura acima ilustra ...
- ... na tabela descrita abaixo.

Corrige-se para

- + [figura no topo ou final da página]
A Figura 2.3 ilustra ...
- + [tabela no topo ou final da página]
... , conforme descrito na Tabela 5.

Se estritamente necessário evitar números:

- + A figura anterior / The previous figure
- + ... na tabela a seguir / ... in the table as follows.



Siglas

SQL, Sql, sql?!

... W3C (*World Wide Web Consortium*) ...



Informalidades

Não

Nope

Jamé

Nunquinha da silva

- ... do algoritmo. E o método ...
- ... do algoritmo. Tipo assim, o método ...
- ... of the algorithm. So, the method ...
- ... the method isn't... can't...

Corrige-se para

- + ... do algoritmo. Ademais, o método ...
- + ... do algoritmo. Por exemplo, o método...
- + ... of the algorithm. Hence, the method ...
- + ... the method is not... cannot



Palavras Vazias

Basicamente / Basically

Obviamente / Obviously

Claramente / Clearly



English in Portuguese

Se existe o termo em português, **use-o** (*performance, accuracy, precision, recall ...*)

Se não existe, use em inglês mas formatado diferente (*throughput*)

Coisa de Brasileiro

Os mano e as mina pira





Data Venia

Data venia, brasileiro
acha que habeas corpus
é sine qua non para
artigo de bona fide.

- ... X takes into consideration Y and Z.
- This happens due to the fact that ...
- Such metric is used to show that X is way superior than Y.
- ... so as to promote a great deal of reading and writing operations per second.

Corrige-se para

- + ... X considers Y and Z.
- + This happens because ...
- + Such metric shows X is better than Y.
- + to support many reading and writing operations per second.



Esse povo é educado demais

Excesso enfraquece a alma e o texto

O objetivo deste artigo **seria**

Nós **tentamos** realizar

Se possível, gostaríamos



It

É necessário

É possível

É interessante ver que

É importante notar que

Está claro que

- but it is necessary to add A and B in order to...
- Therefore, it is not necessary to engage users during the evaluation process.
- Within this segment of agile development, it is clear the importance of ABC.
- One of the problems is that it requires much time to process.
- It is assumed that: [lista de bullets]
- Our solution makes it possible for ABC...
- During this process, it is possible to perform the comparison of design choices...

Corrige-se para

- + ... but it requires adding A and B to ...
- + Therefore, engaging users during the evaluation process is not necessary.
- + ABC is important for agile development.
- + One problem is its time performance.
- + To proceed, we assume that: [lista de bullets]
- + Our solution enables ABC to...
- + This process enables to compare design choices...



Obviamente

O óbvio não é ciência. Evite-o.

Na literatura

Neste artigo

Na área de Computação

Nós por nós e para nós



Nada que não possa ficar pior

The study described in this article shows that X...

>> Our study shows X ...

The method presented in this paper constructs...

>> Our method constructs...

Our investigation presented in this paper seeks
to complement prior research in...

>> Here, our goal is to complement
prior research in ...

Até Agora Linguagem

Texto científico
não é ENEM
Ou historinha
“Como realizei
meu trabalho”

Think in English
Write in English
Avoid **translation**
at all costs

ROTEIRO

1. Processo de Escrita ✓ ✓
2. Estrutura ✓ ✓
3. Linguagem ✓ ✓
4. Estilo
5. Apresentação
6. Finalmente



Dica

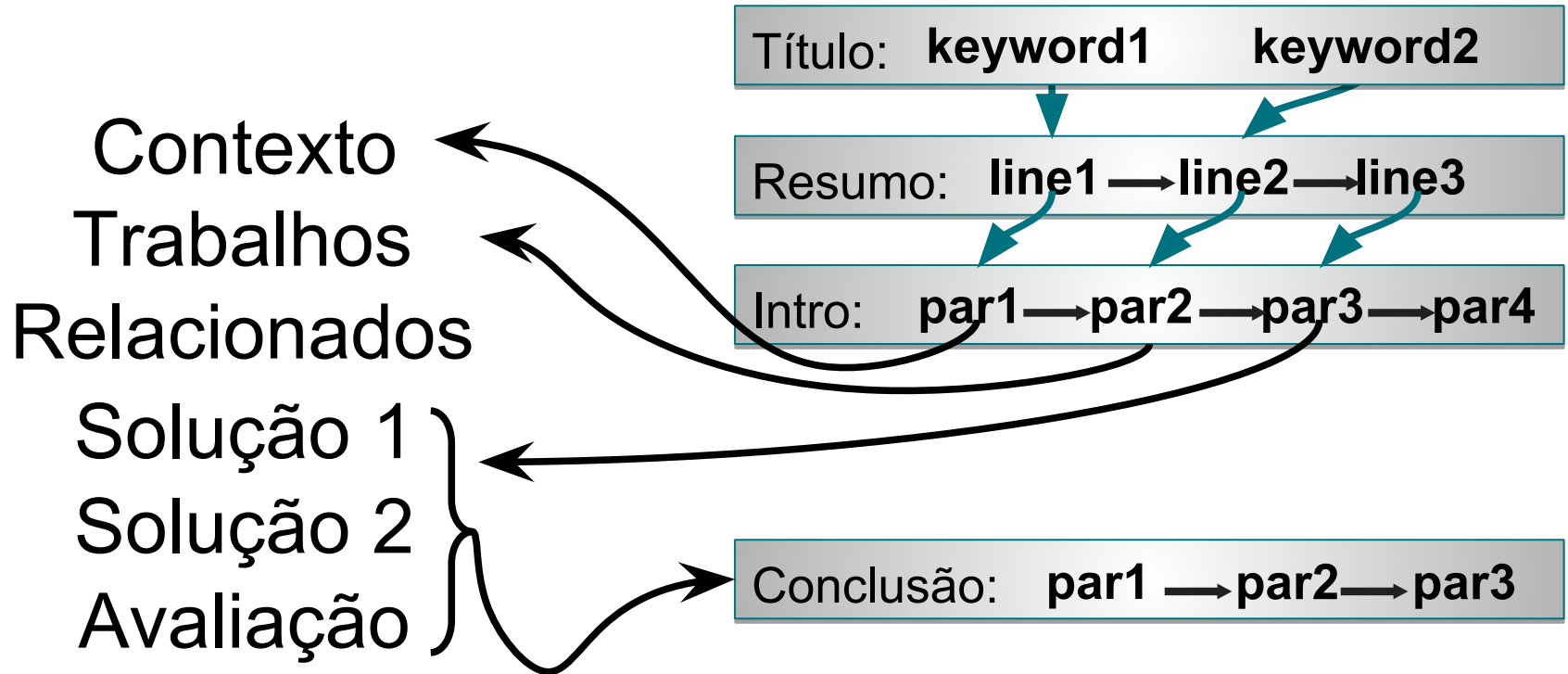
Artigos Científicos
Conferências
Periódicos
Monografia
Dissertação
Tese
Relatório Técnico
Trabalho p/ Disciplina

Estilo

PARTE 4



Fluxo entre componentes = difícil



Estilo e Conteúdo

Artigos científicos possuem regras (ou modelos) de estilo e conteúdo que dependem da área de pesquisa

⇒ verifique artigos recentes publicados na sua área

⇒ tente abstrair um modelo plausível e siga-o

A seguir, algumas regras básicas comuns a vários artigos da Computação (exceções existem)

Aspectos Temporais para Medir a Força da Colaboração no GitHub

Natércia A. Batista, Michele A. Brandão, Ana Paula C. da Silva, Mirella M. Moro

Universidade Federal de Minas Gerais (UFMG) – Belo Horizonte, MG – Brasil

`{natercia,micheleabrandao,ana.coutosilva,mirella}@dcc.ufmg.br`

- Título bacana
- Nomes dos autores
- Universidade (nome em português, sem colocar grupo de pesquisa, departamento, programa de pós, etc., só interessa a instituição)
- Email de contato (um por autor OU por artigo, oficial da instituição)

Artigos formados por seções
Seções por parágrafos
Parágrafos por várias frases

Erros Comuns

- Começar uma seção[§] direto em subseção (deveria começar com texto)
- Seção apenas com uma subseção (deveria ter mais de uma, senão é contínuo)
- Seção composta apenas por lista de itens (artigo não é apresentação de powerpoint)
- Frases soltas no artigo (frases deveriam formar parágrafos)

[§] generalizada como *seção*, mas aplica-se a capítulos e outras unidades de escrita

Artigos formados por seções
Seções por parágrafos
Parágrafos por várias frases

Elementos de Escrita devem ser devidamente referenciados (capítulos, seções, algoritmos, equações, fórmulas, figuras, tabelas, listas, quadros, etc)

Erros Comuns

1. *veja slide sobre Maiúsculas
2. regras aplicáveis a todos os elementos (praticamente)

Fórmulas/equações com números:

✘ This is called triangle inequality and can be resumed in a formula:

$$h(n) \leq c(n, a, n') + h(n') \quad (2)$$

☑ This is called triangle inequality, defined by
*Equation 2.

Figuras, tabelas, etc melhor como sujeito

✘ As we can see in Figure 1 the first level of the division tree would need...

☑ *Figure 1 shows the first level of the division tree ...

Erros Comuns

1. Facilite a leitura do texto e coloque elementos no início ou final da página
2. Exceção para equação e fórmula que são parte integrante do texto

Cuidado com algoritmos:

- ✘ On Algorithm 1, we can see the divide step.
- ✓ Algorithm 1 specifies the division step.

Elementos ficam melhor visualizados se no início ou fim da página

- ✘ The A* search algorithm can be defined as follows.

Algorithm 1: A Search*

```
1 Select the initial set s;  
2 For each ...
```

- ✓ [coloca o algoritmo no início da página e] Algorithm 1 presents the A* search.

“Citações” (*quotes*) geralmente não são utilizadas em Computação
Tem de escrever sua interpretação com suas próprias palavras

Evite citar autores entre “...”

✘ In accordance with Wing [Wing 2006], “computational thinking involves solving problems, designing systems, and understanding human behavior, by drawing on the concepts fundamental to computer science”.

✔ One of the most common definitions for computational thinking is given by Wing [2006]: it uses concepts from computer science to solve problems, design systems and even understand human behavior.

Evite citar autores entre “...”

Neste exemplo, o texto foi editado consideravelmente para deixá-lo mais simples e direto ao ponto que interessa.

✘ Even before robotics be applied in classroom, “Educational theorists such as Papert [Seymour 1980] believed that robotics activities have tremendous potential to improve classroom teaching”. Educators have developed activities to incorporate robotics into the teaching of various subjects, including math, science, and engineering [Benitti 2012].

✔ Robotics may be successfully applied in classroom to teach math, engineering and science, among others [Benitti 2012, Seymor 1980].

Referências são parte importante do artigo
Merecem total atenção

Geralmente

Duas seções que sempre têm referências

1. **Introdução**: referências contextualizam o trabalho bem como reforçam os argumentos de motivação, relevância, atualidade, inovação, justificativa, etc.
2. **Trabalhos Relacionados**: (seção praticamente obrigatória) referências descrevem pesquisas existentes em temas, problemas e soluções semelhantes ao sendo apresentado

Onde referenciar

Próximas ao conteúdo referenciado

- The fast algorithm known so far to find the lowest cost path... [descreve um parágrafo inteiro sobre o algoritmo] ... the best path [Dijkstra 1959]. ⇒ **NÃO!** A referência deveria ter aparecido na primeira frase do parágrafo
- C.E. Shannon and R. M. Fano have developed coding methods ... mas não diz a referência ⇒ **NÃO!** *Precisa* colocar a referência, não pode apenas citar os nomes dos autores.

Como referenciar

Duas vertentes principais

1. A referência como sujeito
 - [Kruskal 1956] presented an algorithm to compute the shortest spanning...
2. Manter os autores como sujeito
 - Kruskal [1956] presented an algorithm to compute the shortest spanning...
 - Brandão and Moro [2017] overview ...

EVITAR:

- [2] overviews methods for coding >>>> Neste caso, inverter a frase: There are methods for coding ... [2].

O que referenciar

Campos obrigatórios no bibtex

PERIÓDICO

@article

- author
- title
- journal
- year
- pages
- volume
- number
- doi

EVENTO CIENTÍFICO

@inproceedings

- author
- title
- booktitle
- year
- address (city, country onde foi realizada)
- pages
- doi

O que referenciar

Campos
obrigatórios
no bibtex

DOUTORADO/

MESTRADO

@phdthesis

@masterthesis

- author
- title
- school
- year
- address (city, country)
- note (url opcional)

RELATÓRIO

TÉCNICO

@techreport

- author
- title
- institution
- year
- number
- address (city, country)
- note (url opcional)

WARNING ao importar bib

- Campos desnecessários devem ser retirados (series, month, etc)
- Nomes invertidos!!! Devem ser escritos na ordem normal
 - booktitle = {Knowledge and Data Mining, International Conference on}
 - journal = {Data and Knowledge Engineering, IEEE Transactions on}
- Nomes errados
 - authors = {Wagner Meira Jr. and Mirella M. Moro} ⇒ Jr., W.M.; Moro, M.M.
 - authors = {Wagner **Meira Jr.** and Mirella M. Moro}
- Address versus location
 - Address = publisher's city & Location = city where the conference took place
 - PPGCC(acm,ieee,etc) ⇐ **Address = city where the conference took place**
- BibTex key: se for muito longa, fica difícil de citar no texto
 - @article{DBLP:journals/jitt/RossettiSZ16, ...
⇒ @article{**RossettiSZ16**, ...

Dicas para o .BIB

- Se espaço for **crítico**, pode-se abreviar o nome das conferências

ORIGINAL: `booktitle = {Proceedings of the
7th International Conference on
Information Technology and Application}`

ABREVIADO:

⇒ `booktitle = {ICITA}`

⇒ `booktitle = {Procs. of ICITA}`

- Ou abrevia todas, ou nenhuma

- Para citar grupo de autores com por exemplo *Moro et al*, tem de alterar o bibtex com uma das duas maneiras:

`author = {Mirella M. Moro and others}`

>> funciona dependendo da classe utilizada

`author = {Mirella M. {Moro et al}}`

- Referências de eventos científicos precisam de local de realização

`address = {City, Country}`

OBS: dblp2.uni-trier.de/db/conf/sigmod tem todas as localizações de realização do SIGMOD (e para a maioria dos eventos em Computação)

Até Agora Estilo

**Estilo, cada um tem o seu.
Ou não.**

Cada área
(ciências ou não)
tem seu estilo.
Dentro da
Computação
também.

ROTEIRO

1. Processo de Escrita ✓ ✓
2. Estrutura ✓ ✓
3. Linguagem ✓ ✓
4. Estilo ✓ ✓
5. Apresentação
6. Finalmente



Dica

Artigos Científicos
Conferências
Periódicos
Monografia
Dissertação
Tese
Relatório Técnico
Trabalho p/ Disciplina

Apresentação

PARTE 5



Roteiro

A. Apresentação?

B. Conteúdo

C. O apresentador

D. Estilo

E. N Exemplos

Apresentação?



Conferência ou Evento

A black and white photograph of a man in a dark suit and light shirt, standing and gesturing with his hands while speaking to an audience. He is positioned on the left side of the frame. In the background, several audience members are seated at tables, some looking towards the speaker. A large video camera is visible on the right side of the frame, pointing towards the speaker. The setting appears to be a lecture hall or conference room with large windows in the background.

Defesa



$$\begin{aligned} & \sum_{n=1}^{\infty} \int_0^1 \frac{(1-x)^{2n}}{n!} dx = \sum_{n=1}^{\infty} \frac{(1-x)^{2n}}{n!} \Big|_0^1 \quad (1) \\ & = \sum_{n=1}^{\infty} \frac{(1-x)^{2n}}{n!} \Big|_0^1 \quad \text{memorável!} \\ & \text{Se } f(x) = \int_0^1 \frac{(1-x)^{2n}}{n!} dx = \frac{1}{n!} \quad (\text{Laplace 472}) \end{aligned}$$

$$\begin{aligned} & \sum_{n=1}^{\infty} \int_0^1 \frac{(1-x)^{2n}}{n!} dx = \sum_{n=1}^{\infty} \frac{(1-x)^{2n}}{n!} \Big|_0^1 \quad (1) \\ & = \sum_{n=1}^{\infty} \frac{(1-x)^{2n}}{n!} \Big|_0^1 \quad \text{memorável!} \\ & \text{Se } f(x) = \int_0^1 \frac{(1-x)^{2n}}{n!} dx = \frac{1}{n!} \quad (\text{Laplace 472}) \end{aligned}$$

Sala de Aula



Potenciais Investidores



Reunião de Projeto

A black and white photograph of two women in a computer lab. The woman on the left has dark hair and is resting her chin on her hand, looking at a computer monitor. The woman on the right has long blonde hair and is looking directly at the camera. In the background, another person is visible at a desk with a computer. The text 'Sua Orientadora' is overlaid in white on the bottom left.

Sua Orientadora

**Apresentação
de Trabalho
Científico e afins
para quem quer
que seja**

Qual o objetivo?



**Objetivo é sempre
vender o peixe**

**Contribuição central
do trabalho científico**

Pode ser também

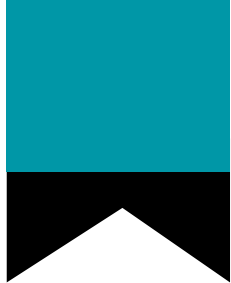
- Defesa
- O apresentador
- O grupo de pesquisa
- A instituição
- ...



**Em uma conferência,
o objetivo pode ser
incentivar a ler o artigo 😊**

Por onde
começar?

START



B. Conteúdo

O que colocar na apresentação

Como organizar seu conteúdo

Em que ordem

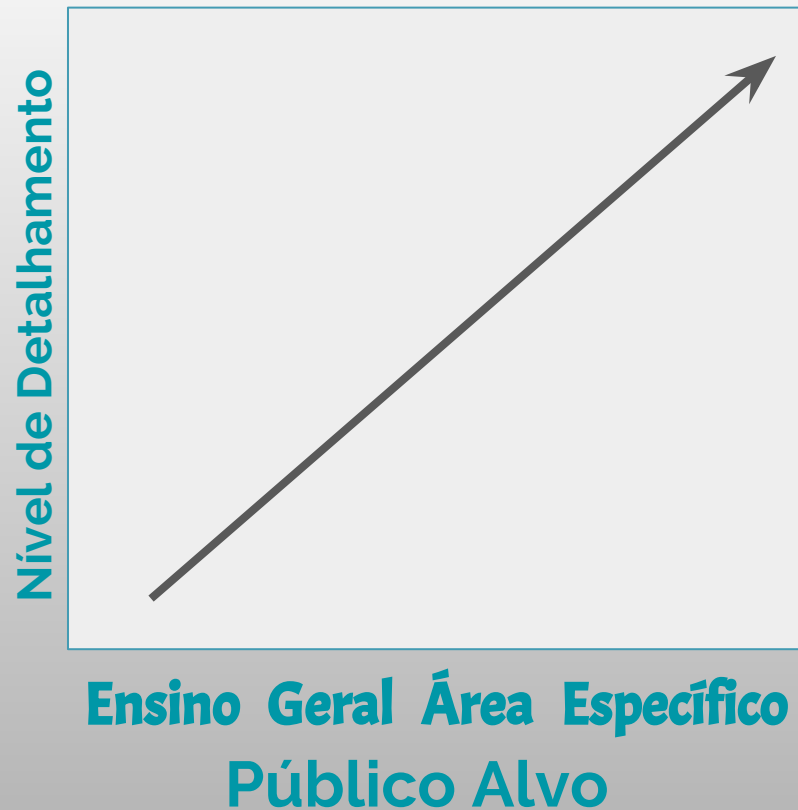
Princípio Simples:
os detalhes estão
no artigo

O que é ou não detalhe?

Planejar é preciso

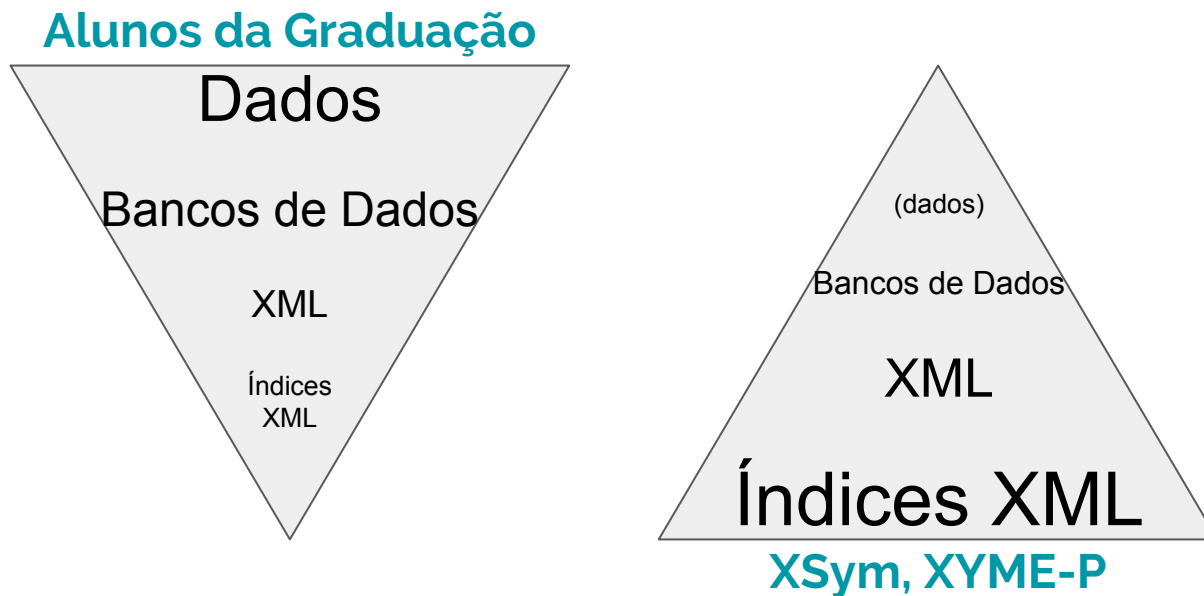
Detalhes ou Não

Conteúdo



Por exemplo

“XHI – Índice Híbrido para Dados XML”



Conteúdo

Contexto

Problema

Solução

Avaliação





Contexto e Problema

O estado-da-arte

Objetivo: localizar o seu trabalho

2 Estratégias para Contexto

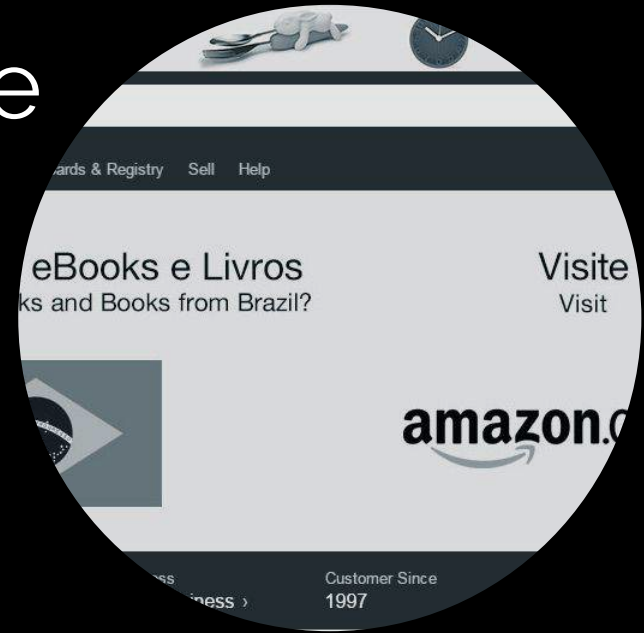
- (1) motivar através de exemplos
- (2) apresentar o estado-da-arte
 - [Contribuição: localizada no contexto]

#1 Contexto = Motivar através de Exemplos

A seguir: palestra ERBD 2017

**A. Se você sabe o
que quer ...**

1. Entre no navegador
2. Endereço da loja online
3. Digite o título
4. Pressione <ENTER>



Ler notícias
Website do jornal
favorito



**B. Se você não
sabe o que quer**

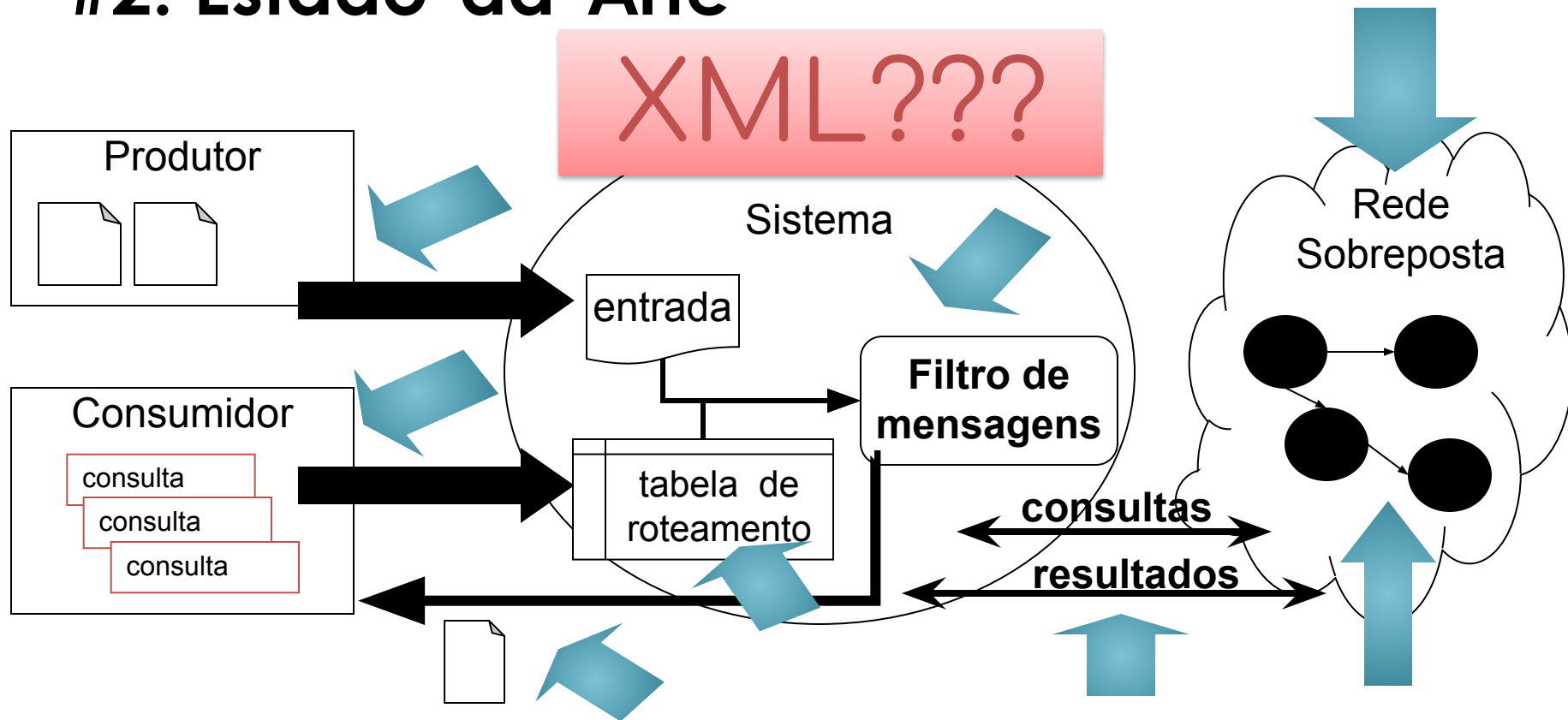


Procurar
um livro
seria
assim?

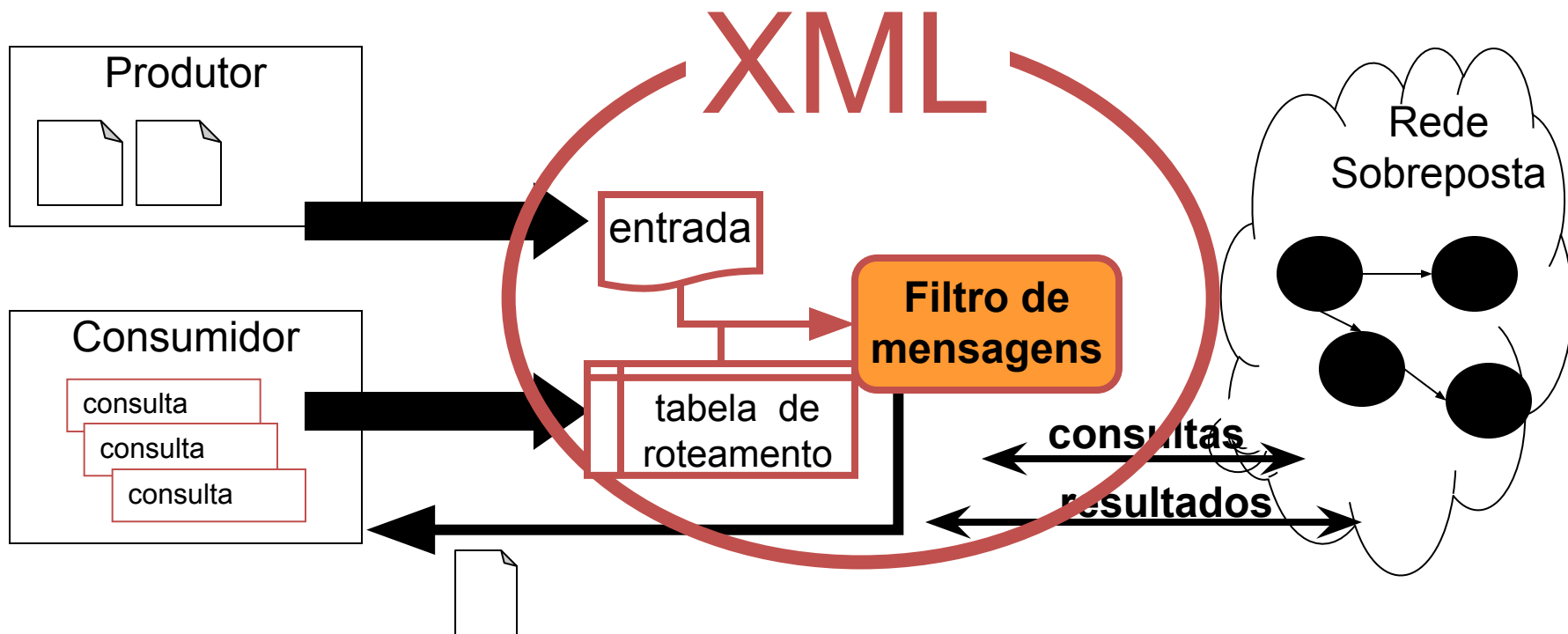


**Na Web?
Seria cair aí**

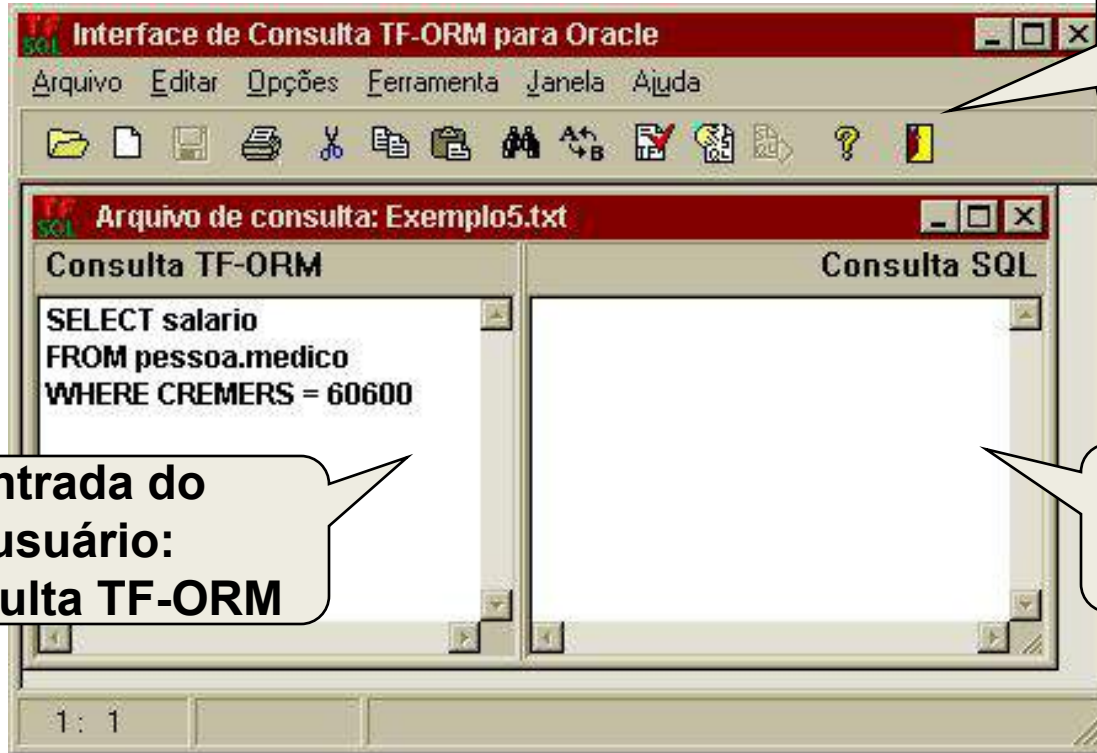
#2: Estado-da-Arte



Contexto e Problema



Problema e Solução: Demo



Verificar, mapear e
executar consulta
TF-ORM

Entrada do
usuário:
Consulta TF-ORM

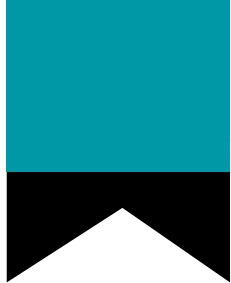
Resultado do
mapeamento:
Consulta SQL

Conteúdo:

Segue o do artigo

Porém, não a forma

Contexto, Problema, Solução, Avaliação (exemplos a seguir)



C. O/A Apresentador/a

A apresentação é composta pelo apresentador e pelo conjunto de slides

A seguir, SETE erros

#1. Monotonia Monotom

Solução: treinar
antes com uma
plateia (mesmo
que pequena)



#2. Falta de Ensaio

Solução: saber o que
dizer em cada slide;
decorar os primeiros
5 minutos



#3. Ticks

Solução: treinar
com plateia que
alguém vai
apontá-los 😊



#4. Relógio

Solução: não o levar,
considerar outro tipo
de controle de
tempo (celular
discretamente na
bancada)



#5. Exagero de Humor e Gestos

Solução:
controlar-se



#6. Fala Rápida (atropelando tudo)

Solução: treinar
antes com uma
plateia

**Seguir a fórmula:
1 minuto por slide**



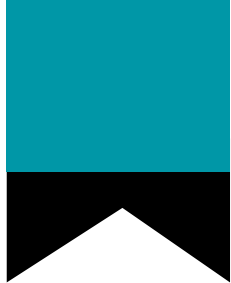
#7. Falta de Contato Visual

Solução: manter-se
de frente para
plateia sempre;
focar em alguém
conhecido



Grandes Apresentadores





D. Estilo

Como na moda, estilo é tudo

A seguir, SETE erros

(exemplos reais de estudantes das minhas turmas)

Chemical Compound Dataset

#1.
Frases
Compleatas

- The Chemical Compound Dataset can be retrieved through the URL:
http://oldwww.comlab.ox.ac.uk/oucl/groups/mac_hlearn/PTE
- The dataset contains 340 chemical compounds, 24 different atoms, 66 atom types, and 4 types of bonds.
- On average 27 vertices and 28 edges per graph.
- The largest one contains 214 edges and 214 vertices. So the discovered patterns are much like tree, though they do contains some cycles.
- The goal is to find the common chemical compound substructure

#1. Frases Completas

Mais sutil que o
anterior, mas ainda
assim...

Related Work

Depth-First and breadth-first search exhaust all possible paths

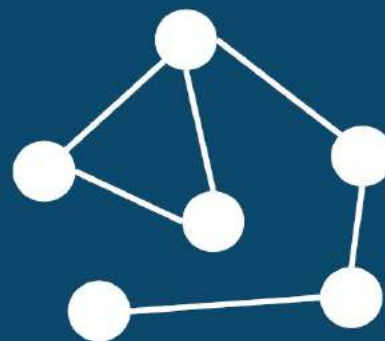
Bellman-Ford algorithm compute shortest paths from one vertex to all of the others handling edges with negative weights

Dijkstra's algorithm is the best so far to find the shortest path given a source and a goal

These approaches are not goal oriented and they explore unnecessary paths to find the best one

#1. Frases Completas

Às vezes,
tudo bem
MAS vejam que
não está poluído



Graph

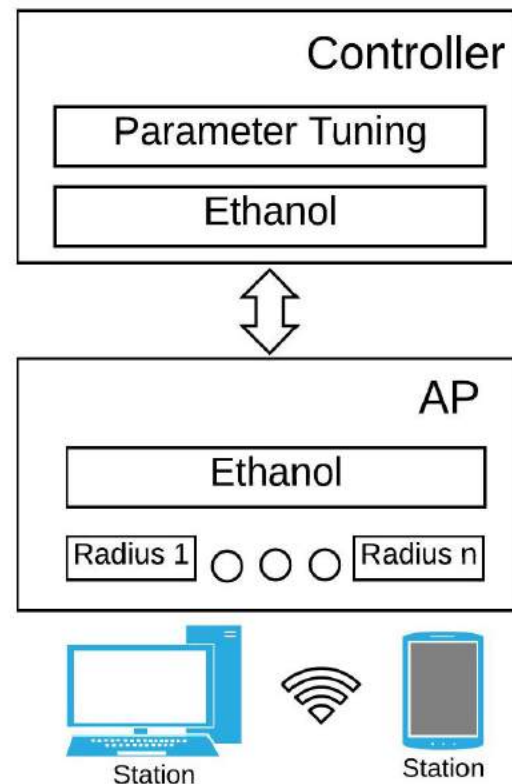
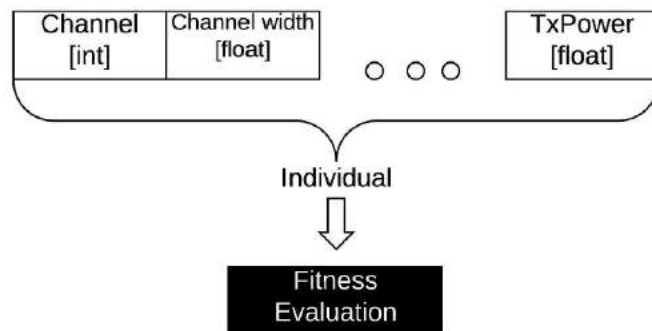
An ordered pair $G = (\mathbf{V}, \mathbf{E})$ comprising a set $\mathbf{V}$ of vertices or nodes with a set $\mathbf{E}$ of edges.

#2. Slide Poluído

Muita info no
mesmo slide

This work: summarized

- Parameter tuning module uses a GA to optimize wi-fi parameters for the Access Point, with the goal of optimizing users' QoE



#2. Slide Poluído

Muita info no
mesmo slide

Introduction

PCA

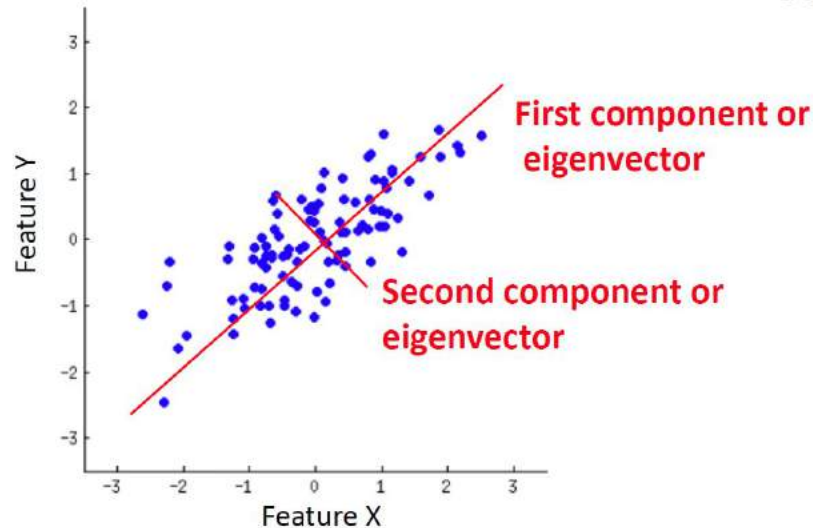
Eigenpictures

Eigenfaces

Experimental
Results

Conclusions

Principal Components Analysis (PCA)



$$\text{Cov}(X, Y) = \frac{1}{n} \sum_{i=1}^n (x_i - m_x)(y_i - m_y)$$

$$\begin{pmatrix} \text{Cov}(X, X), & \text{Cov}(X, Y) \\ \text{Cov}(Y, X), & \text{Cov}(Y, Y) \end{pmatrix}$$

#3. Fontes pequenas ou difíceis

Fake news **Introduction**

The problem is very hard to solve

Because we can't solve it efficiently ourselves.

Solution must be unbiased

Because fake news will exist at both ends of a subjective spectrum.

Untrue news is not necessarily fake

A source might publish a fact that simply is not known to be false.

#3. Fontes pequenas ou difíceis

SELECT - Return all books under a certain price (2/2)

Query

```
PREFIX dc: <http://purl.org/dc/elements/1.1/> .  
PREFIX : <http://example.org/book/> .  
PREFIX ns: <http://example.org/ns#> .
```

```
SELECT ?book ?title
```

```
WHERE
```

```
{ ?book dc:title ?title .  
  ?book ns:price ?price . FILTER ( ?price < 40 )  
}
```

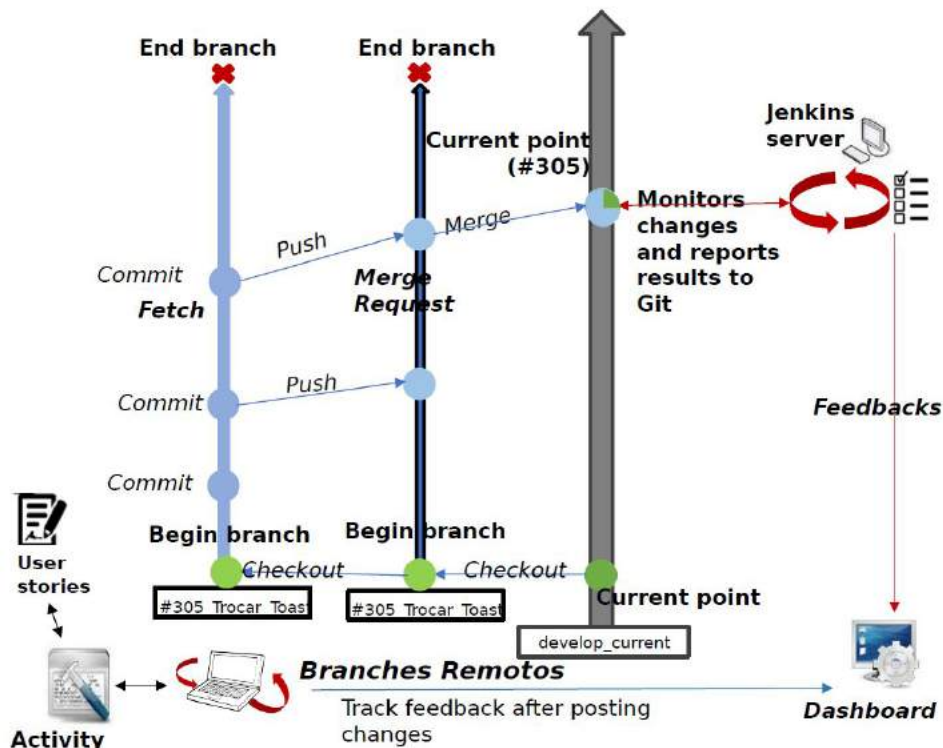
Result

book	title
:book2	"The Semantic Web"



#3. Fontes pequenas ou difíceis

Methodology



#2 + #3
Slide
poluído +
Fonte
pequena

Fake news
Related work

Many smart people are actively working on it.

Fake news challenges offers a USD10.000 prize for anyone who can provide an efficient solution to the problem.

Facebook tried to include the source on news posts but it did not turn out to be that efficient.



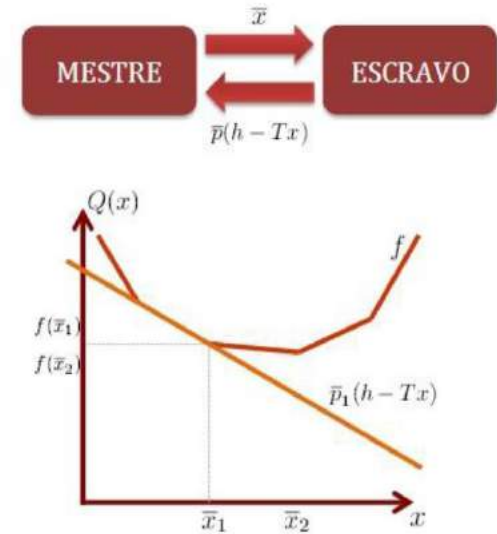
#4. Images ilegíveis

(+ #2 poluído)
(+ #3 fonte ilegível)

Método de Decomposição de Benders

Note que:

- Como o problema mestre é um problema aproximado, sua solução ótima é sempre um **limite inferior** (LB) para o valor da solução real
 - Tal problema é dito **Mestre Relaxado** possui somente parte das restrições (ditas cortes) que definem a região viável do problema escravo.
- Se eu avalio a solução do mestre no problema original, eu obtenho um **limite superior** (UB)
- Quando o algoritmo se encerra?
 - Quando $UB = LB$
 - Na prática: Quando $UB - LB < \epsilon$



Experiments and Performance Study

#4.
Images
ilegíveis

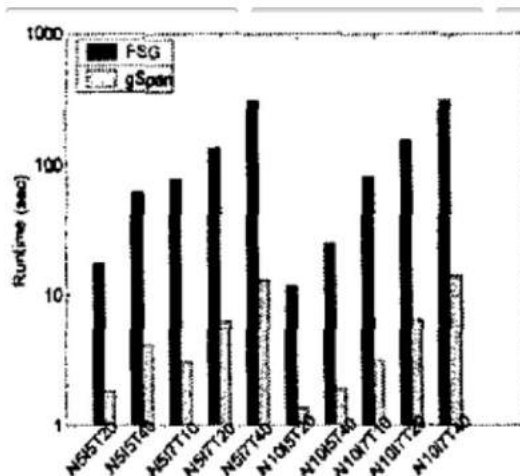


Figure 4. Runtime: Synthetic data

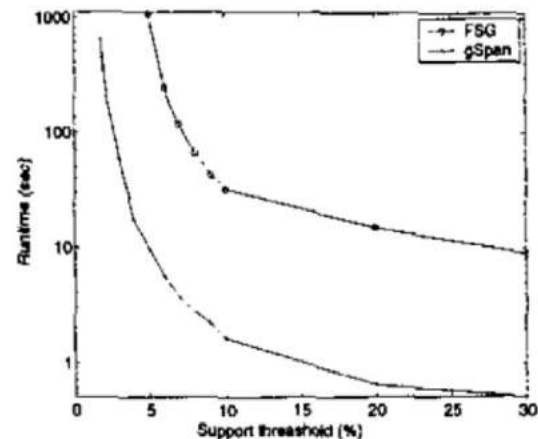


Figure 5. Runtime: Chemical data

#5. Animações Malucas

Aquelas que as letras das
palavras das frases
chegam voando do
nada... ou girando...

#6.
Eros
Otográficos

Percisa ezemplo?

#7.
Falta

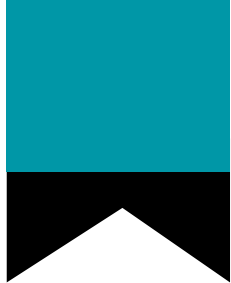
de

Fluxo

SUGESTÕES DE CORREÇÃO

1. Frases completas
2. Slides poluídos
3. Fonte pequenas ou Fontes difíceis
4. Imagens ilegíveis
5. Animações malucas
6. Eros otográficos
7. Falta de fluxo

1. Palavras-chave
2. 1 slide = 1 informação
3. Fonte tamanho ≥ 20 Arial, sans-serif
4. Novas imagens
5. Animações necessárias
6. Revisão
7. Revisão COM ensaio



E. N Exemplos

É a melhor maneira de mostrar o que fazer e o que evitar

Exemplos reais de apresentações minhas

Exemplo: muita informação + fonte pequena

- The widespread employment of XML motivates the development of efficient methods for manipulating XML data.
- Query languages (e.g. XQuery, XPath) take into consideration the structure of the data and enable querying both on its structure and simple values.
- Tree-pattern queries are considered among the most important and yet more complex queries to process.
- Holistic processing techniques have outperformed more conventional solutions.
- Many techniques have been proposed to handle tree-pattern queries holistically (e.g. TwigStack, XML filtering, ViST, PRIX).
- Index structures (e.g. B+-tree, XR-tree, XB-tree) have been introduced to further improve performance.
- A common characteristic for all holistic approaches is that some preprocessing is required, either on the data (e.g. TwigStack) or both the data and query (e.g. ViST, PRIX).
- Thus the need for a “dedicated” (native or relational) XML storage manager where all data resides and can be preprocessed.
- Problem: While many processing methods have already been proposed for such queries (and were shown to have promising performance and potential), none of them has found its way to any of the existing lightweight XML engines (e.g. Galax XQuery processor, eXist).

Exemplo: muita informação+fonte pequena

- Solução #1: dividir slide
- 1 slide = 1 informação

- The widespread employment of XML motivates the development of efficient methods for manipulating XML data.
- Query languages (e.g. XQuery, XPath) take into consideration the structure of the data and enable querying both on its structure and simple values.
- Tree-pattern queries are considered among the most important and yet more complex queries to process.
- Holistic processing techniques have outperformed more conventional solutions.
- Many techniques have been proposed to handle tree-pattern queries holistically (e.g. TwigStack, XML filtering, ViST, PRiX).
- Index structures (e.g. B+-tree, XR-tree, XB-tree) have been introduced to further improve performance.
- A common characteristic for all holistic approaches is that some preprocessing is required, either on the data (e.g. TwigStack) or both the data and query (e.g. ViST, PRiX).
- Thus the need for a “dedicated” (native or relational) XML storage manager where all data resides and can be preprocessed.
- Problem: While many processing methods have already been proposed for such queries (and were shown to have promising performance and potential), none of them has found its way to any of the existing lightweight XML engines (e.g. Galax XQuery processor, eXist).

XML everywhere
XQuery, XPath
structure + values
Tree-pattern queries

Holistic Tree Pattern Query

- Better than conventional solutions
- E.g.: TwigStack, filtering, ViST, PRiX
- Index: B+-tree, XR-tree, XB-tree
- Preprocessing required
 - Data
 - Data + query

Problems

- Need for a “dedicated” XML storage manager (native or relational)
- No algorithm on lightweight XML engines

Exemplo: muita informação+fonte pequena

- Solução #2: usar cores
- 1 informação = 1 cor

```
SELECT t2.nome, t3.salario
FROM PESSOA t0, MEDICO t1,
     PESSOA_NOME t2, MEDICO_SALARIO t3
WHERE t0.oid = t1.oid AND t0.oid = t2.oid AND
     t1.rid = t3.rid AND t3.salario > 2000 AND
     t2.t_timei <= '01/01/2000 AND
     t3.t_timei <= '01/01/2000'
```

MAPEAMENTO:

Tabelas

Relacionamentos

Restrições dados

Restrições tempo

Exemplo: importância da cor em exagero

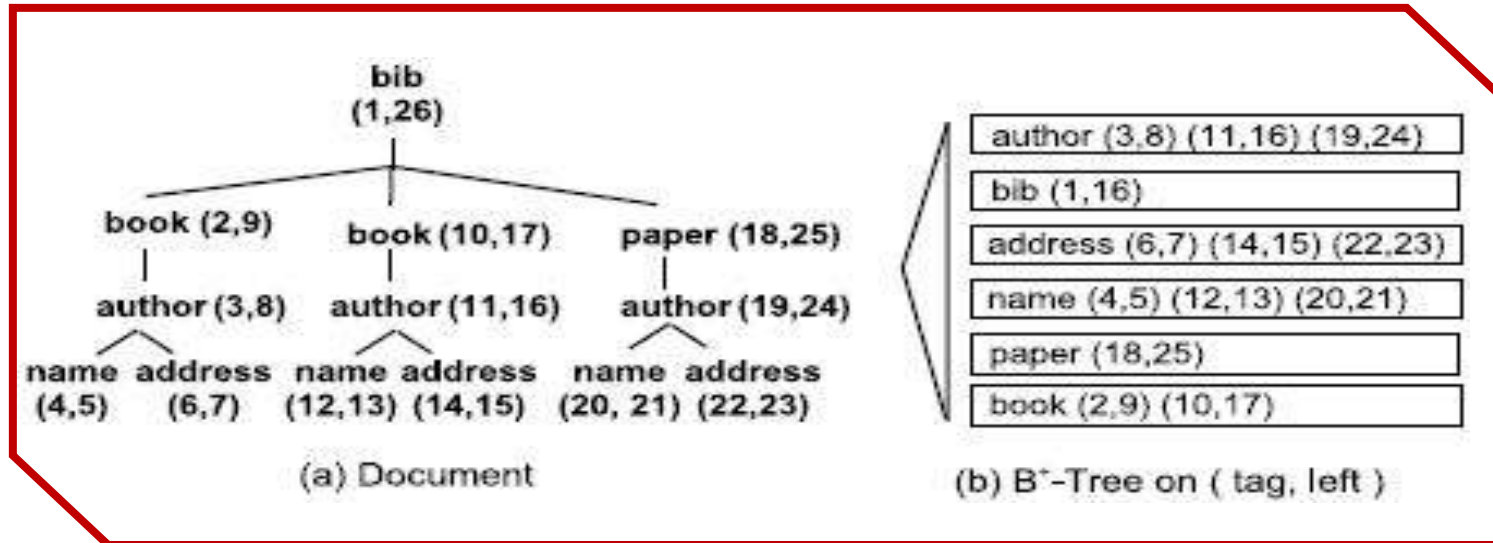
```
SELECT t2.nome, t3.salario
FROM PESSOA t0, MEDICO t1,
     PESSOA_NOME t2, MEDICO_SALARIO t3
WHERE t0.oid = t1.oid AND t0.oid = t2.oid AND
     t1.rid = t3.rid AND t3.salario > 2000 AND
     t2.t_timei <= '01/01/2000' AND
     t3.t_timei <= '01/01/2000'
```

```
SELECT t2.nome, t3.salario
FROM PESSOA t0, MEDICO t1,
     PESSOA_NOME t2, MEDICO_SALARIO t3
WHERE t0.oid = t1.oid AND t0.oid = t2.oid AND
     t1.rid = t3.rid AND t3.salario > 2000 AND
     t2.t_timei <= '01/01/2000' AND
     t3.t_timei <= '01/01/2000'
```

Exemplo: copy/paste imagem

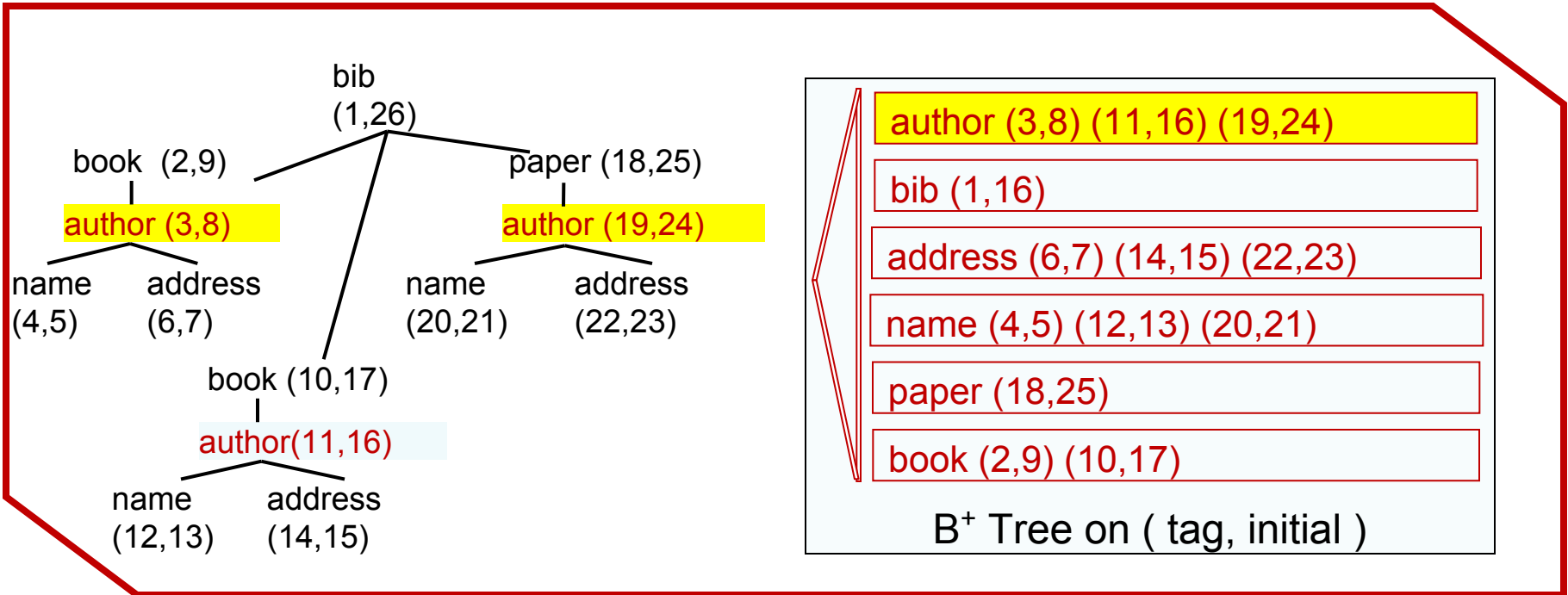
Copy/paste imagem do artigo

- Nem sempre legível
- Horrível, informação demais



Exemplo: copy/paste imagem

Solução: refazer imagem usando software de apresentação



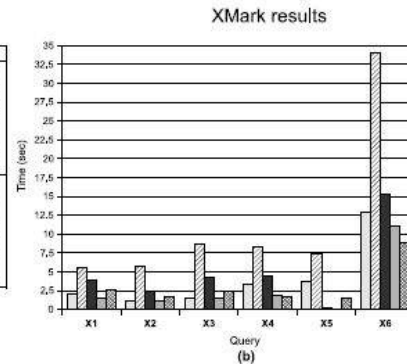
Exemplo: copy/paste imagem

Solução: refazer imagem usando software de apresentação

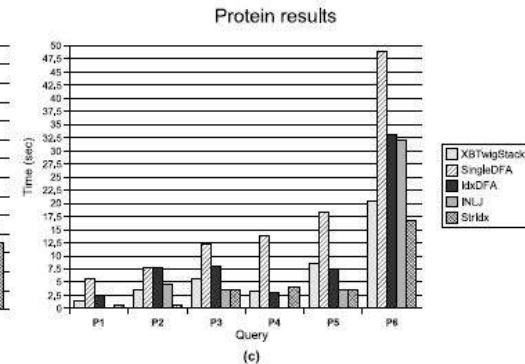
	Twig	Single	IdxDFA	INLJ	StrIdx	Sel(%)	#Nodes
X1	100	100	100	140	100	100	600002
X2	100	100	100	119,93	79,95	78,95	636944
X3	99,66	100	100	114,47	61,37	80	891372
X4	100	100	78,45	42,03	42,14	30,02	1055612
X5	100	100	2,73	4,7	40,06	2,7	1364519
X6	85,82	100	73,35	135,4	78,89	60,35	3099331

Percentage of nodes accessed per method,
query selectivity, number of nodes in document

(a)



(b)



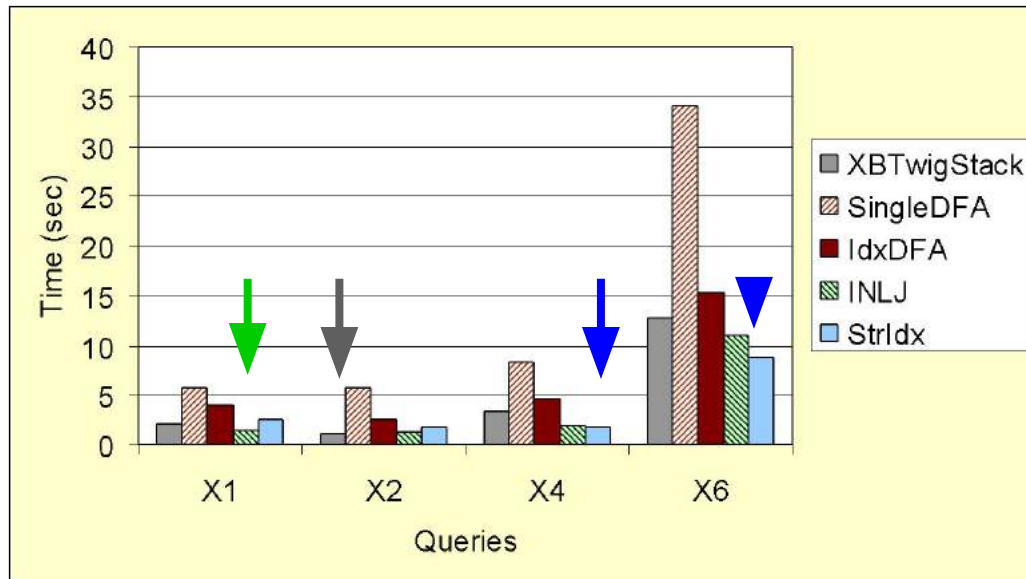
(c)

Figure 6: Results for XMark and Protein datasets

- Resultados XMARK
 - Não tem classificação fixa
 - SingleDFA tem pior desempenho
 - StrIdx melhor ou comparável a melhor solução

Exemplo: copy/paste imagem

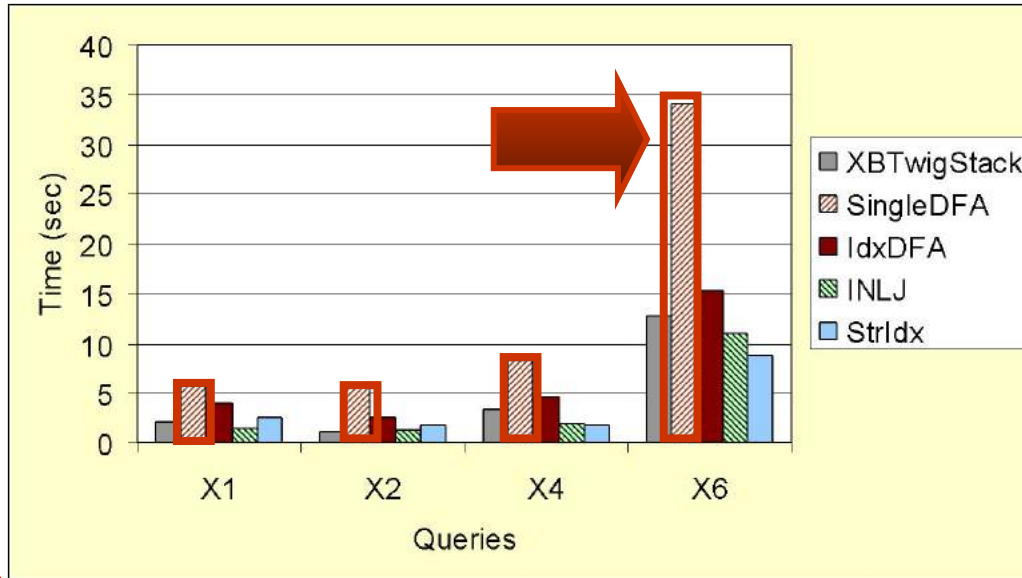
Solução: refazer imagem usando software de apresentação
+ animação (necessária)



- Resultados XMARK: 1. não tem classificação fixa

Exemplo: copy/paste imagem

Solução: refazer imagem usando software de apresentação
+ animação (necessária)



- Resultados XMARK: 2 SingleDFA pior desempenho

Começando o Curso

Atualize seus *bookmarks*

- **BDBCOMP**
 - <http://www.lbd.dcc.ufmg.br/bdbcomp/>
- **Portal Periódicos/CAPES**
 - <http://www.periodicos.capes.gov.br>
- **DBLP** – biblioteca digital Ciência da Computação
 - <http://www.informatik.uni-trier.de/~ley/db>
- **ACM Digital Library** – *Assoc. for Computing Machinery*
 - <http://portal.acm.org/dl.cfm>
- **IEEE Xplore** – *Inst. of Electrical and Electronics Eng.*
 - <http://ieeexplore.ieee.org>
- **CiteSeer** – biblioteca digital de literatura científica
 - <http://citeseer.nj.nec.com/cs>
- **Science Direct** – journals publicados pela *Elsevier*
 - <http://www.sciencedirect.com/science/journals/computerscience>
- **Google** <http://www.google.com>

Começando o Curso

Atualize seus *bookmarks*

- **BDBCOMP**

<http://www.lbd.dcc.ufmg.br/bdbcomp>

- **DBLP**

<http://www.informatik.uni-trier.de/~ley/db>

- **Portal Periódicos/CAPES**

<http://www.periodicos.capes.gov.br>

- **ACM Digital Library** – *Assoc. for Computing Machinery*

<http://portal.acm.org>

- **IEEE Xplore** – *Inst. of Electrical and Electronics Eng.*

<http://ieeexplore.ieee.org>

- **SciVerse / Science Direct** – journals publicados pela *Elsevier*

<http://www.sciencedirect.com/science/journals/computerscience>

web

Como se faz Pesquisa

1) Tema em Aberto

LEIA LEIA LEIA LEIA

Mas ler o que mesmo?

- Tuuuuuuuuuuuudo! (cientificamente falando)
- Inicie pelos trabalhos recentes do seu orientador (facilita a tarefa de conhecer o seu trabalho)
- A partir desses trabalhos, selecione algumas referências bibliográficas e leia também
- Leia os *proceedings* das conferências mais importantes da sua área dos últimos três anos
 - SBBD, VLDB, ICDE, SIGMOD
- Aluno de doutorado : um artigo por dia

Ler é a ÚNICA solução

Orientador

- Seus artigos recentes
- Dissertações/Teses recentes

Relacionados

- A partir das referências
- A partir de buscas

Eventos Periódicos

- Principais (VLDB, SIGMOD, SBBD)
- De 2018, 2017, ...

Mais exemplos

Reais de outras pessoas

Com *muitos* erros ao mesmo tempo

Logo do projeto em todos os slides

Menu com navegação

Fonte pequena

Muito espaço em branco

Data



Principles

Data Mining

Career

Data Mining

Graph Mining

Structural
Correlation
Patterns

Knowledge
Transfer

Web
Observatory

Analytics
Systems

Summary

- Research education is a long process.
- Every work should present clear contributions.
- Cooperation is a key research mechanism.
- Target real problems, beyond CS, whenever possible.
- Research results should become technology and be transferred.

Como eu faria

Lembrando: LaTeX
não foi feito para
produzir slides, e
sim **texto**

Principles

Research Education

- Long process

Every work

- Clear contribution

Cooperation

- Key research mechanism

Real problems

- Beyond CS, whenever possible

Research Results

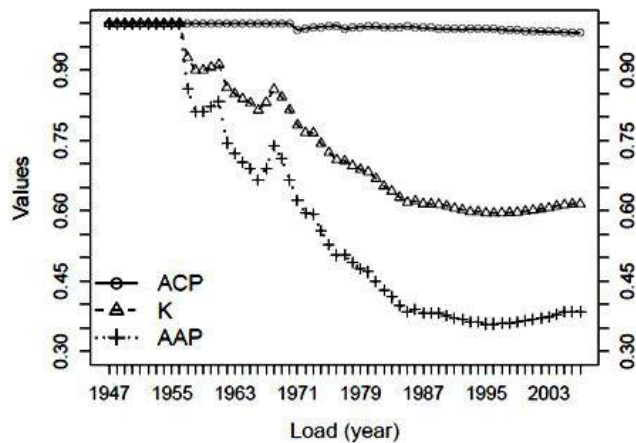
- Transferred technology

The screenshot shows a presentation slide titled "Principles". On the left is a blue sidebar menu with the "w inweb" logo at the top and a list of topics: Data Mining, Graph Mining, Structural Correlation Patterns, Knowledge Transfer, Web Observatory, Analytics Systems, and Summary. The "Principles" title is in the top right. The main content area lists five principles, each preceded by a blue square bullet point:

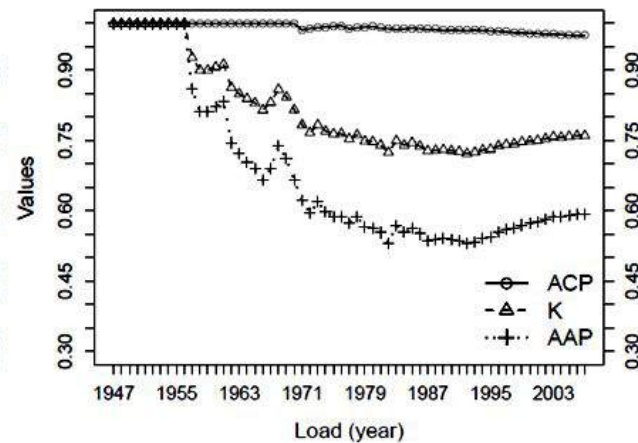
- Research education is a long process.
- Every work should present clear contributions.
- Cooperation is a key research mechanism.
- Target real problems, beyond CS, whenever possible.
- Research results should become technology and be transferred.

At the bottom of the slide, it says "Oct 3, 2015" on the left and "Main" on the right.

LaTeX de novo
 Monte de info extra
 Gráficos ilegíveis
 Preto e branco?!
 Com legenda



(a) INDi



(b) ALL-ALL

Figure : KISTI collection

ALL-ALL

Outperforms INDi around 16,6%, under the K metric.

Letra verde em
fundo preto (1980?)

Gráfico ilegível

Legenda
copy-paste do
artigo

Explicações
coloridas em frases
completas

Nome do evento +
data

Gini Coefficient

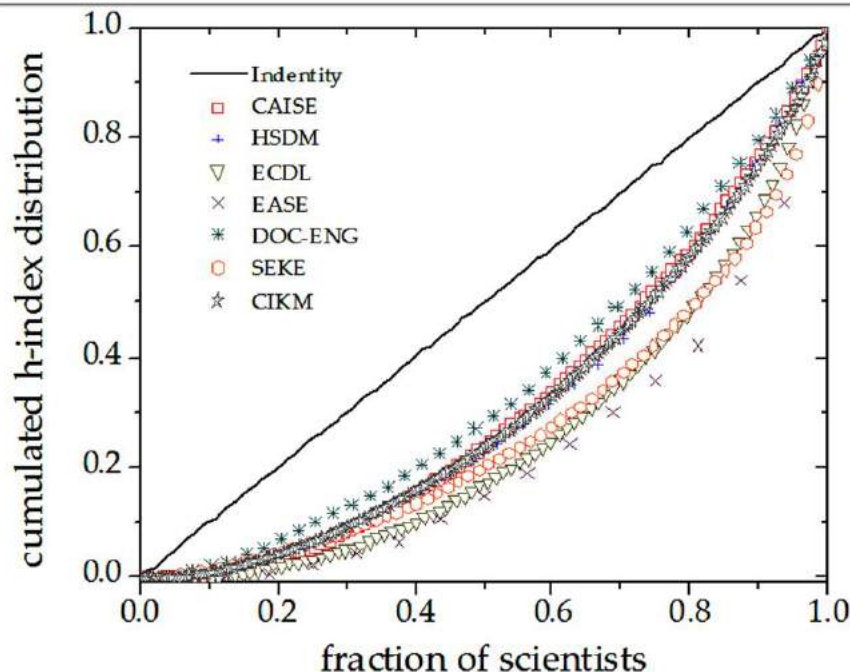


Fig. 2. Lorenz curves for the h-index distributions of researches in conferences of Engineering Software.

The Lorenz curve is a graph that represents the cumulative distribution of a probability density function.

The Gini coefficient is calculated as the area between the *perfect equality* line and the observed Lorenz curve.

This coefficient is directly proportional to the inequality of the distribution.

A low value of Gini coefficient indicates a more equal distribution among the parts, and a high value indicates a more unequal distribution.

Frases completas

Mais um uso
exagerado de cores

Falta fluxo

Fontes distintas

Logo

“Slide”

SQL (1/3)

SQL is a standardized language used to manage a Relational Database [2].

SQL consists of a **data definition language (DDL)**, a **data manipulation language (DML)** and a **data query language (DQL)**.

SQL-DDL is used to define and modify the schema of a database.

SQL-DML is used to modify the records of the database.

SQL-DQL is the query language, based on relational algebra, used to access data stored in the database.



Frases completas

Logo

Espaço em branco

Problema e Motivação

- ▶ Mas, seria possível olhar pra dentro desta caixa e capturar as suas informações internas a ela?
- ▶ Se isto fosse possível, como relacionaríamos este tipo de informação com a especificação e a execução do workflow?
- ▶ Seria possível relacionar estes dados sem a realização de um pós-processamento?

Frases completas

;

Template esquisito

Título do slide à
direita na vertical

The use of collected knowledge can be effectively used towards pathfinding improvement;

We implemented a fast algorithm for pathfinding;

In the worst case our approach performs same as the state-of-art.

Conclusion

Você tem 15 minutos para apresentar

Vai perder um slide inteiro *lendo* o conteúdo?

Sendo que em conferência, todos apresentam isso?

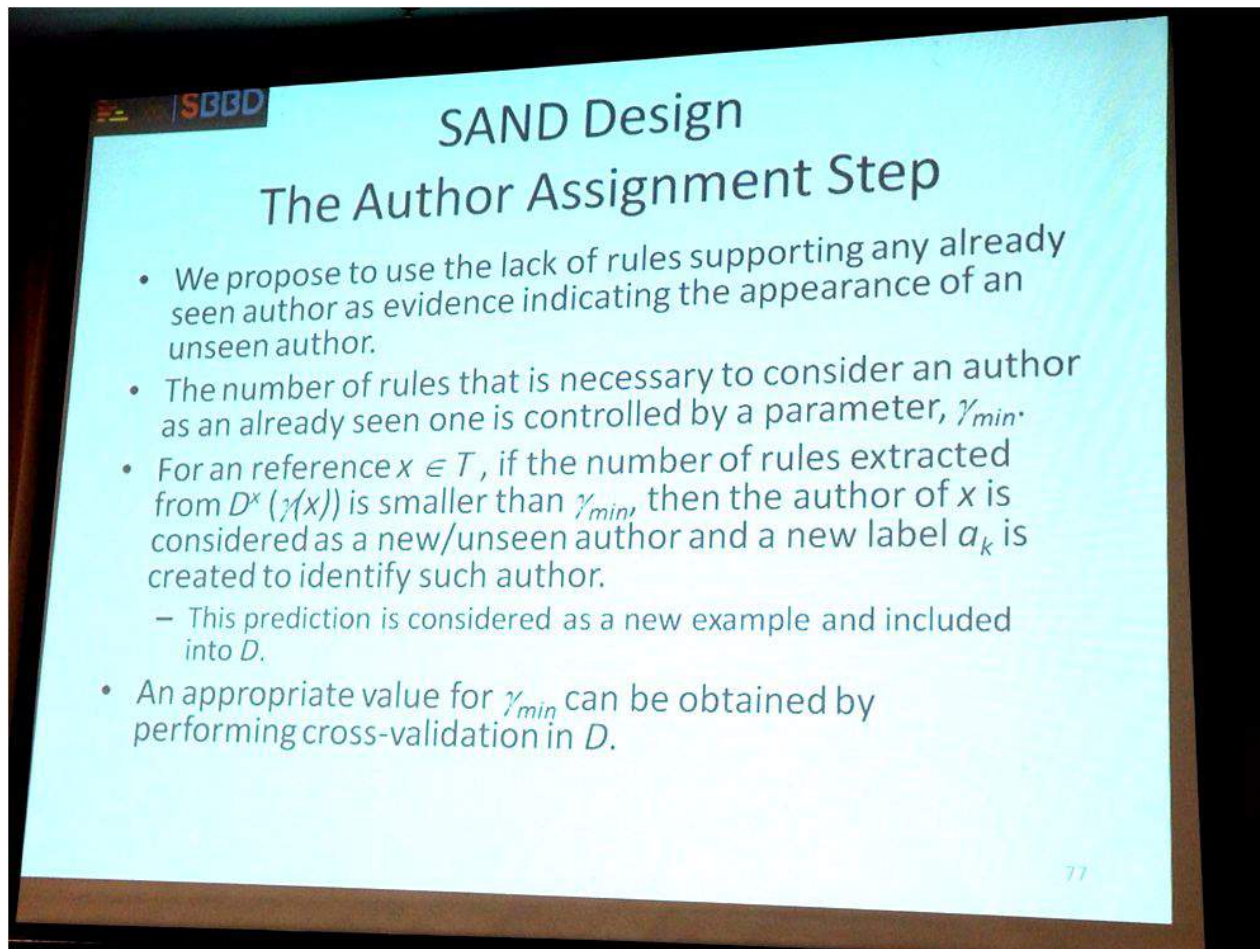
E . ????

Outline

- Introduction.
- Similarity join.
- The similarity *Wide-join* operator.
- Experiments.
- Conclusion.

(foto)

Definições formais
ESTÃO no artigo e
lá devem ficar



SAND Design

The Author Assignment Step

- We propose to use the lack of rules supporting any already seen author as evidence indicating the appearance of an unseen author.
- The number of rules that is necessary to consider an author as an already seen one is controlled by a parameter, γ_{min} .
- For an reference $x \in T$, if the number of rules extracted from $D^x(\gamma(x))$ is smaller than γ_{min} , then the author of x is considered as a new/unseen author and a new label a_k is created to identify such author.
 - This prediction is considered as a new example and included into D .
- An appropriate value for γ_{min} can be obtained by performing cross-validation in D .

77

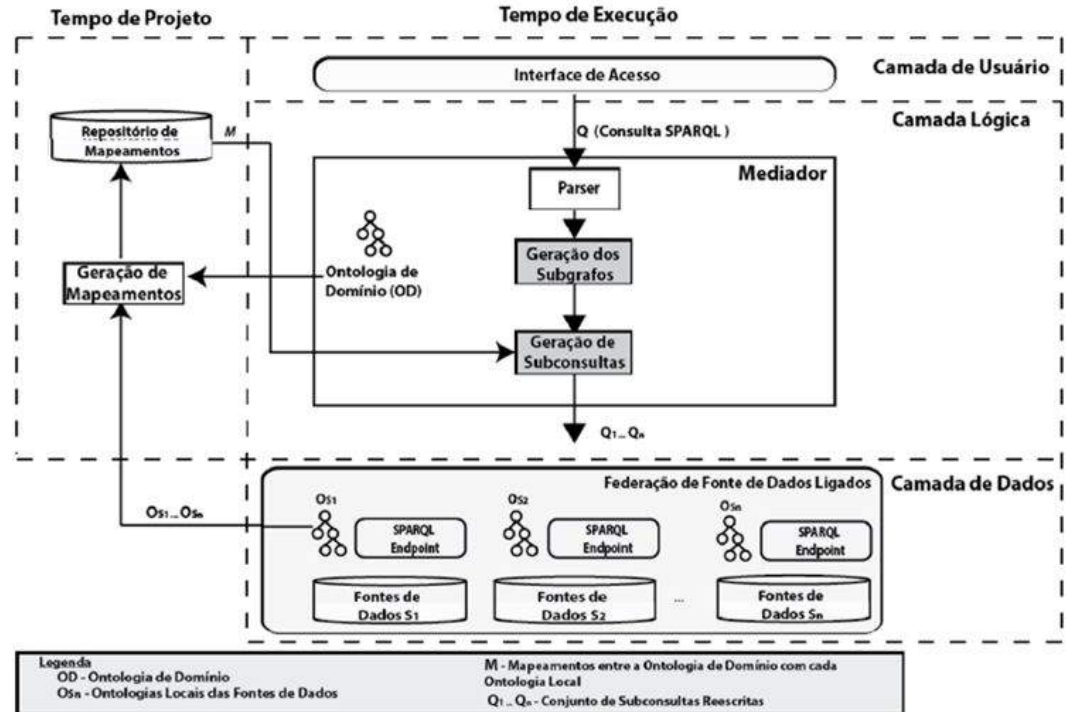
- oLinDa: *Query Decomposition over Linked Data Federation*

É a imagem mais importante do artigo e da apresentação

Ilegível!

Fácil de refazê-la com o software apropriado

Para que o bullet se é só um item?



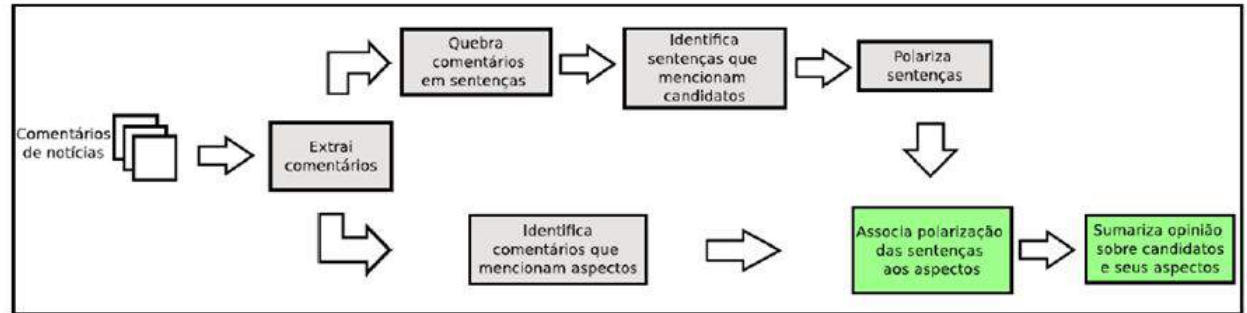
Mineração da opinião sobre aspectos de candidatos a eleições em comentários de notícias

É a imagem mais
importante do artigo
e da apresentação

Ilegível!

Fácil de refazê-la
com o software
apropriado

Logo da instituição
em todos os slides



Processo para mineração das opiniões em nível de aspecto

Chameleon

Melhor que isso, só colocando o logo como rodapé E como fundo (#sqn)

- Otimização de sistemas MapReduce tornou-se comum;
- Sistemas com algum tipo de otimização ganham sucesso no mercado;
- Análise do histórico de execução auxilia na otimização de novas tarefas;
- Utilizando ajuste automático de parâmetros, o Chameleon auxilia na otimização de consultas do Hive em tempo real.

(foto)

Lembre sempre:
“detalhes no artigo”

Experimental Evaluation Metrics

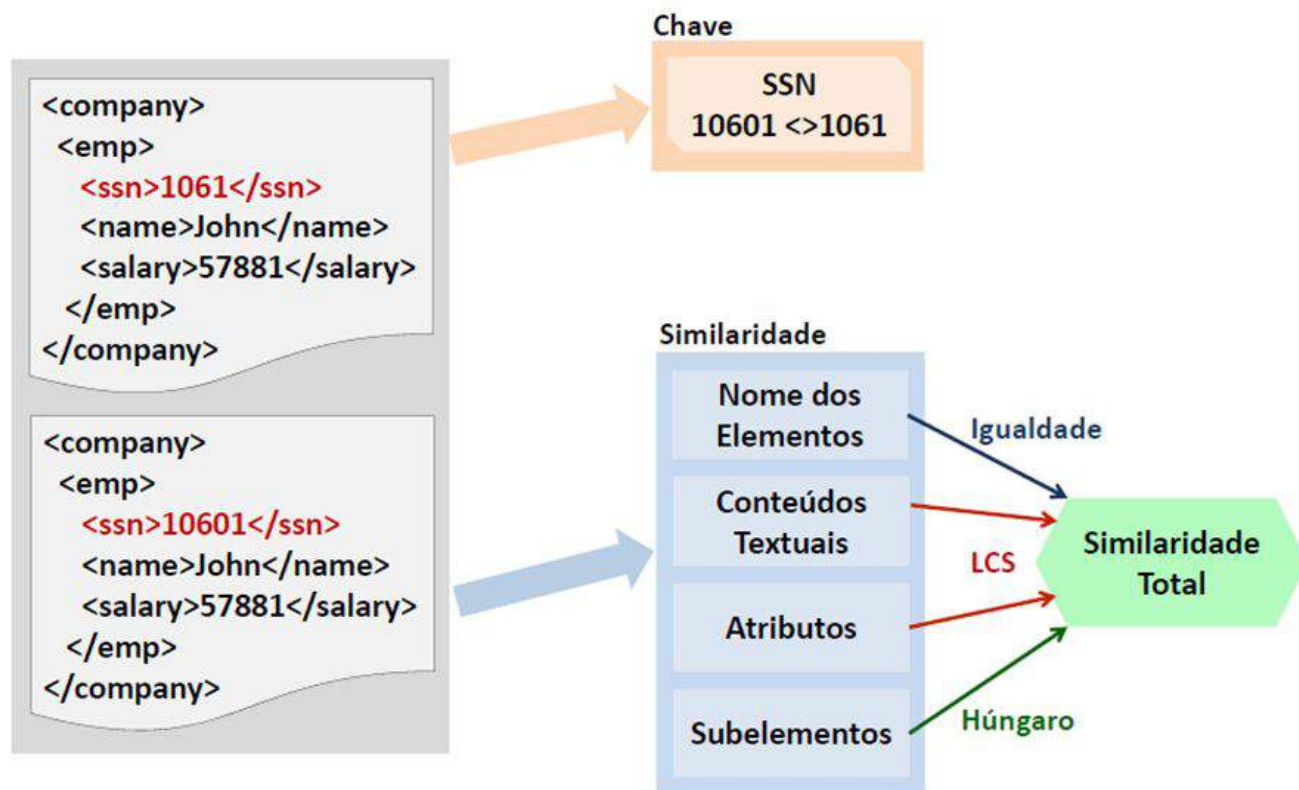
Metrics

$$ACP = \frac{1}{N} \sum_{i=1}^e \sum_{j=1}^t \frac{n_{ij}^2}{n_i} \quad (1)$$
$$AAP = \frac{1}{N} \sum_{j=1}^t \sum_{i=1}^e \frac{n_{ij}^2}{n_j} \quad (2)$$
$$K = \sqrt{ACP \times AAP} \quad (3)$$

Reducing Fragmentation in Incremental Auth... SBBD 23 / 46

Os Bons Exemplos

Refez figura
Clean
Claro
Rodapé mínimo



Recomendação de objetos

**Duo de slides
perfeito**

1: definição curta

- * Sugerir objetos adequados
- * Para a pessoa certa
- * Na hora certa
- * No contexto certo

**Duo de slides
perfeito**

2: exemplo claro

Professor Palazzo @
SBBD 2014

No futebol

Onde está o gol?



Fred



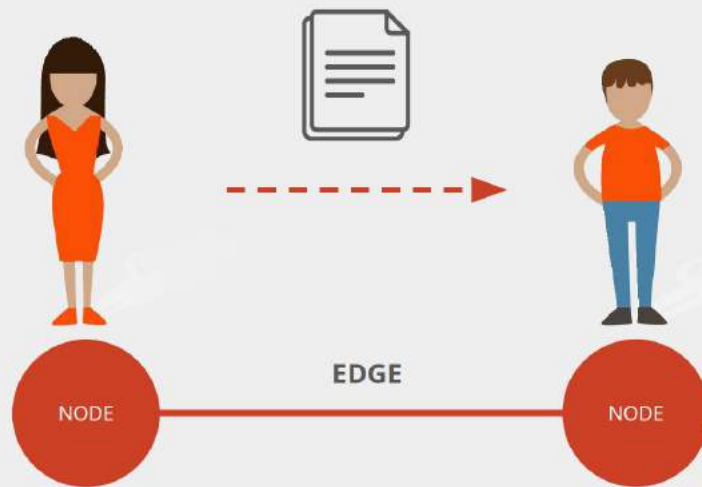
Criativo
Inteligente
Claro
Objetivo

NEIGHBORHOOD OVERLAP



Claro
Objetivo
Para uma
definição
complexa

CO-AUTHORSHIP SOCIAL NETWORKS



Claro
Objetivo
Metodologia
com animação

METHODOLOGY



Contexto claro
Objetivo
Cores não
tradicionais

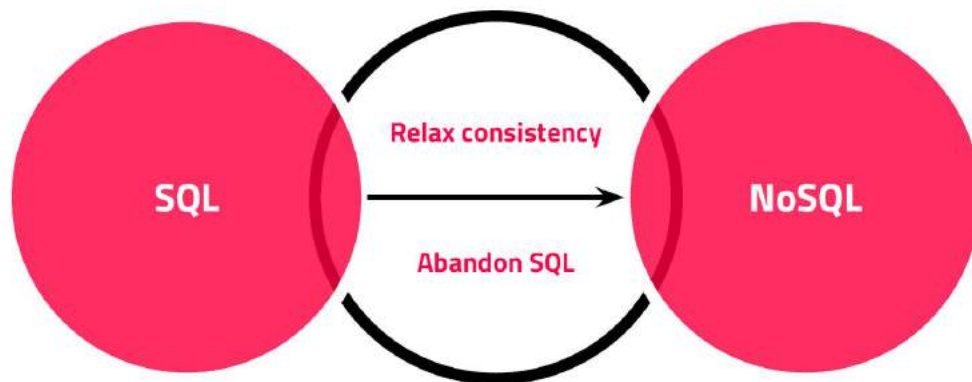


DATABASES

- Used in almost all applications
- Dominance of relational databases
- **But what is the current scenario of data?**

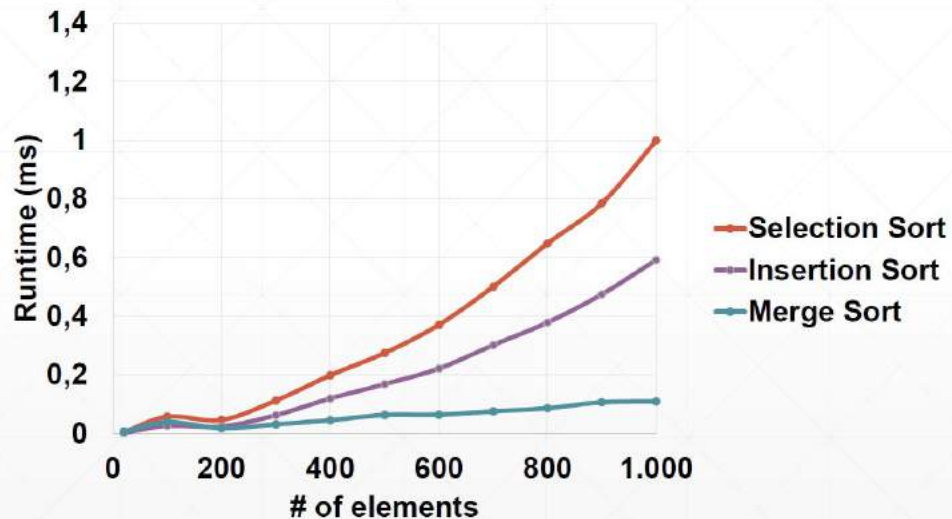
Claro
Objetivo
Cores não
tradicionais

SQL x NoSQL



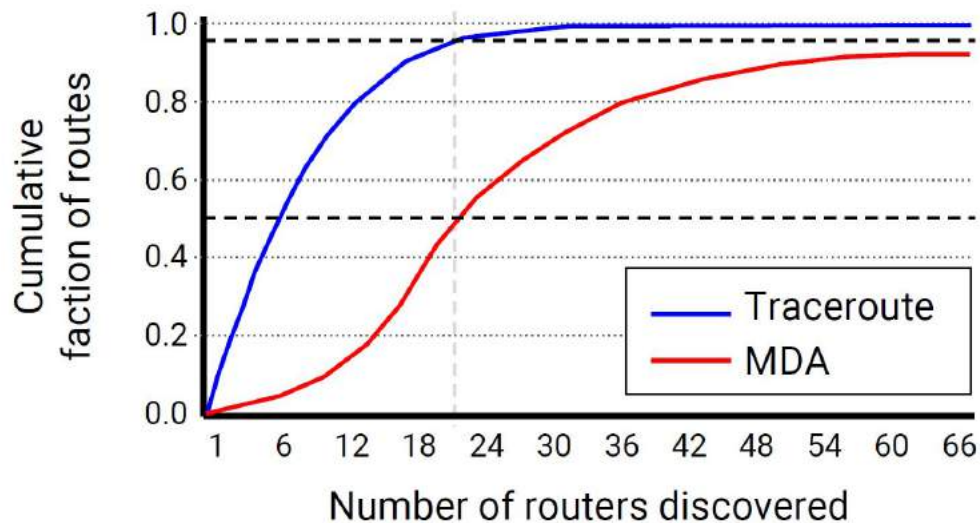
Refez o gráfico
do artigo
Legenda e eixos
legíveis
Cores

Running time



MDA vs Traceroute

Refez o gráfico
Texto Claro
Cores
Linhas tracejadas
(animação)



Momento de
transição na
apresentação

A photograph showing a business meeting. Two people are seated at a desk with laptops. They are looking at a clipboard that has several sheets of paper with charts and graphs on it. One person is pointing at a chart with their finger, and the other is holding a pen over the same chart. The charts include pie charts and line graphs. The text 'EXPERIMENTAL RESULTS' is overlaid on the image in a white, bold, sans-serif font. There is a small red horizontal line above the word 'EXPERIMENTAL'.

**EXPERIMENTAL
RESULTS**

**Momento de
transição na
apresentação**

**Enfatiza o
problema
existente**

No method focusing on
communicability found in
chatbots' literature

**Momento de
transição na
apresentação**

**Enfatiza o
objetivo do
trabalho**



Our goal

Explore the interaction of aspects of the HCI curriculum with students across different institutions, disciplines and countries.

**Apresentação:
clara e objetiva**

Finalmente

PARTE 6

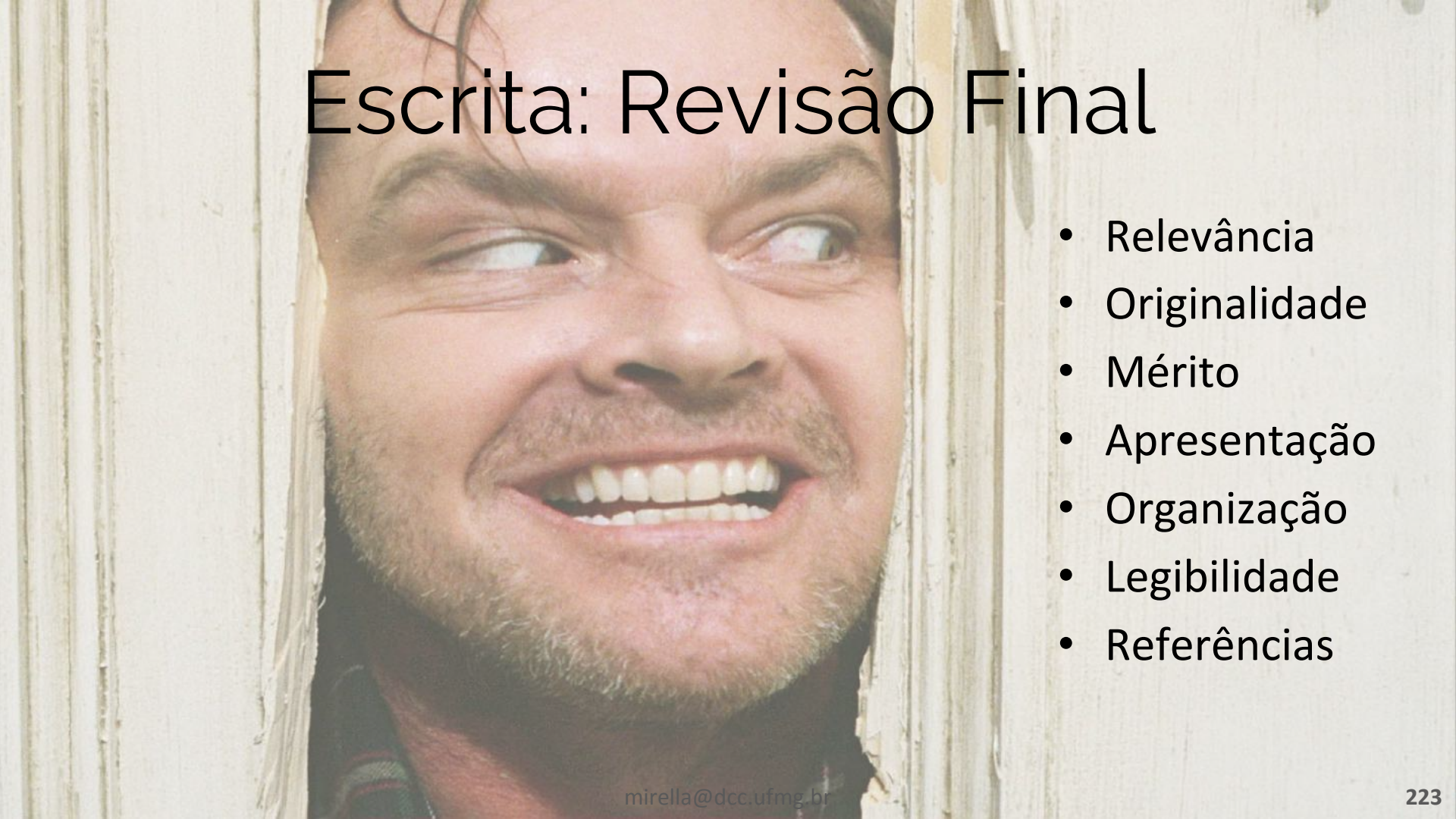


ESCRITA: REVISÃO FINAL

Verificar antes da Submissão

SE TODOS AUTORES LERAM O ARTIGO INTEIRO E ESTÃO DE ACORDO

- Ortografia de título, nomes dos autores e filiação
- Imprima o artigo (no formato final de submissão): tudo legível
- Tenha certeza absoluta da data e do horário limites para submissão de trabalhos

A close-up photograph of Jack Nicholson's face as he peeks through a jagged hole in a light-colored wooden wall. He has a wide, toothy grin and his eyes are looking upwards and to the sides with a mischievous expression. The lighting is soft, highlighting his facial features and the texture of the wood.

Escrita: Revisão Final

- Relevância
- Originalidade
- Mérito
- Apresentação
- Organização
- Legibilidade
- Referências

Escrita Científica

1. Título
2. Autor (es)
3. Resumo
4. Introdução
5. Corpo
6. Conclusão
7. Referências

1. Comece com algum, melhore
2. Você, colaboradores, orientador/a
3. Contexto, problema, solução, avaliação
4. Contexto + motivação + problema + estado-da-arte + contribuições + avaliação + organização
5. Estado-da-arte: comparação com o proposto
Contribuições: trabalho desenvolvido
Validação: análise, estudo, experimentos
6. Estado atual + próximos passos
7. Local (IES), nacional, internacional

Apresentação

Criatividade para motivar plateia

Detalhes estão no texto

Cuidado com 7+7 erros

Saber decor 5 primeiros slides

Treinar com quem não conhece o trabalho

ESCRITA E APRESENTAÇÃO DE TRABALHOS CIENTÍFICOS



www.dcc.ufmg.br/~mirella →
mirella@dcc.ufmg.br

*Como condensar artigos
Modelo de dissertação/tese do PPGCC*