



# Gerência de Dados Semi- Estruturados



Vanessa Braganholo  
vanessa@ic.uff.br

# Apresentações

---

- ▶ Um pouco mais sobre mim
  - ▶ Vanessa Braganholo
  - ▶ [www.ic.uff.br/~vanessa](http://www.ic.uff.br/~vanessa)
  
- ▶ E vocês?
  - ▶ Nome?
  - ▶ Área de pesquisa? / Período da graduação?
  - ▶ Emprego? Estágio? Projeto de Aplicação? Iniciação Científica?
  - ▶ Expectativas em relação a esta disciplina?

# Dados Semi-Estruturados

---

- ▶ Dados irregulares, incompletos, sem necessariamente estarem de acordo com um esquema
- ▶ Seu representante mais famoso: XML

# Forma de Avaliação

---

- ▶ Média =  
(Prova + Participação + 2 x Trabalho +  
Avaliação de Artigos) / 5
- ▶ **APROVADO:** (Presença  $\geq 75\%$ ) E (Média  $\geq 6$ )
- ▶ **VS:** (Aluno de Graduação) E (Presença  $\geq 75\%$ ) E ( $4 \leq$   
Média  $< 6$ )
- ▶ Será aprovado na VS se tirar nota maior ou igual a 6.

# Grupos

---

- ▶ Todas as atividades, exceto a prova e avaliação de artigos, são em grupo
  - ▶ Mas as notas são individuais!
- ▶ Deve ser o mesmo durante todo o curso
  - ▶ 3 participantes para graduação
  - ▶ 2 participantes para mestrado
  - ▶ Individual para doutorado
- ▶ Definir na primeira semana de aula
  - ▶ Enviar por e-mail (assunto: GDSE - Grupo) a matrícula e o nome completo dos participantes

# Dinâmica do curso

---

- ▶ Aulas convencionais
- ▶ Leitura de artigos
  - ▶ Entrega de resumos (todos os grupos)
  - ▶ Apresentações de artigos (1 grupo por artigo)
- ▶ Desenvolvimento do trabalho
  - ▶ Apresentação dos trabalhos será dividida em duas partes
    - ▶ Andamento
    - ▶ Resultado Final

# Tópicos das Leituras

---

- ▶ Introdução a Dados Semi-Estruturados
- ▶ Futuro da pesquisa em Banco de Dados
- ▶ Visões XML
- ▶ Integração de esquemas XML
- ▶ Processamento de consultas distribuídas
- ▶ Arquitetura de SGBD XML nativo

# Resumos

---

## ▶ Características

- ▶ Sucintos, não passando de 1 página A4 (fonte 12, margem de 2 cm)

## ▶ Conteúdo

- ▶ Título do artigo e nomes dos membros do grupo
- ▶ Resumo em si

## ▶ Importante

- ▶ Focar no que realmente interessa no artigo
- ▶ Ter princípio, meio e fim (evitar o uso de *bullets*)
- ▶ Encerrar com 1 parágrafo que emita a sua opinião sobre o artigo (pontos positivos e negativos do artigo)

## ▶ Atraso: multa de 1 ponto por dia

## ▶ Nota dos resumos entra na nota de Participação, que vale 20% da nota final da disciplina

# Apresentações de Artigo

---

- ▶ Cada grupo ficará encarregado de apresentar um dos artigos da semana
  - ▶ Apresentação de 30 minutos
  - ▶ Uso de projetor
- ▶ Os demais grupos ficarão encarregados de defender ou criticar os artigos
- ▶ Nota da apresentação de artigos entra na nota de Participação, que vale 20% da nota final da disciplina

# Trabalho

---

- ▶ **Objetivo:**
  - ▶ Estudo ou aplicação de técnicas avançadas de gerência de dados semi-estruturados
- ▶ Tema deve ser definido nas primeiras semanas de aula

# Trabalho

---

- ▶ **Tipos de trabalho**
  - ▶ Teórico: foco maior na descrição dos trabalhos relacionados
  - ▶ Implementação: foco maior na descrição da ferramenta e exemplo de uso
- ▶ **Resultado do trabalho:**
  - ▶ Artigo no formato da SBC (8 a 10 páginas)
  - ▶ Apresentações de andamento e final
- ▶ **Conteúdo do artigo**
  - ▶ A motivação e o objetivo do trabalho
  - ▶ Trabalhos relacionados
  - ▶ Resultados obtidos
  - ▶ Considerações finais

# Avaliação de Artigos

---

- ▶ Os trabalhos serão submetidos via um sistema, simulando uma mini-conferência
- ▶ Mas, como funciona uma conferência de verdade?

# Processo de Avaliação de Artigos

---



# Processo de Avaliação de Artigos

---



# Processo de Avaliação de Artigos

---

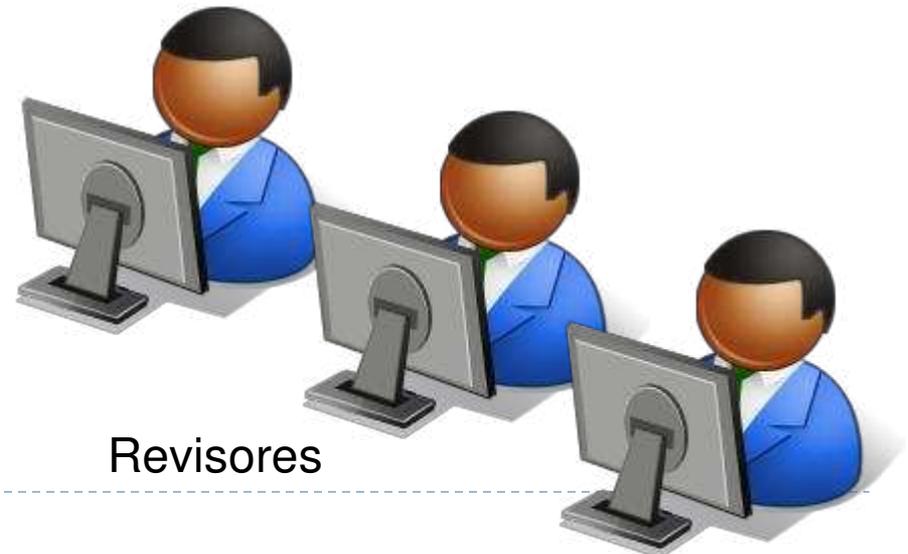


# Processo de Avaliação de Artigos

---



3 Avaliações



# Processo de Avaliação de Artigos

---



# Avaliação de Artigos

---

- ▶ Cada aluno estará cadastrado como membro do comitê de programa dessa mini-conferência e receberá artigos para avaliar
  - ▶ A avaliação dos artigos vale 20% da Média (item Avaliação de Artigos)
  - ▶ Ao final, todos receberão anonimamente as avaliações dos seus artigos
  - ▶ A avaliação feita pelos colegas não afetará a nota do artigo

# Apresentações do Trabalho

---

- ▶ **1ª Parte**

- ▶ Contexto do trabalho
- ▶ Objetivo
- ▶ Andamento atual

- ▶ **2ª Parte**

- ▶ Apresentação final do trabalho
- ▶ Resultados obtidos
- ▶ Relato de experiência

# Sites para busca de artigos...

---

- ▶ <http://scholar.google.com.br>
- ▶ <http://www.informatik.uni-trier.de/~ley/db>
- ▶ <http://www.scopus.com>
- ▶ <http://ieeexplore.ieee.org>
- ▶ <http://portal.acm.org>
- ▶ <http://citeseer.ist.psu.edu>
  
- ▶ Usem uma ferramenta para controlar as suas referências: <http://www.zotero.org>

# Página do curso



www.ic.uff.br/~vanessa/courses/2011.2/gdse.html

Vanessa Braganholo Location: // Courses / 2011.2 / GDSE

**Gerência de Dados Semi-Estruturados**

Horário: quartas e sextas, de 9:00 às 11:00

Local: Sala 234-B, bloco E (quartas) e Sala 230-A, bloco D (sextas)

Lista: <http://groups.google.com.br/group/gdse-uff-2011-2>  
(importante: todos os alunos **devem** se cadastrar nessa lista)

**Avaliação**

Média =  
(Prova + Participação + 2 x Trabalho + Avaliação de Artigos) / 5

**APROVADO:**  
(Presença  $\geq$  75%) E (Média  $\geq$  6)

**VERIFICAÇÃO SUPLEMENTAR:**  
(Aluno de Graduação) E (Presença  $\geq$  75%) E (4  $\leq$  Média < 6)

Será aprovado na VS o aluno que tirar nota maior ou igual a 6.

**REPROVADO:** caso contrário

Leiam as regras do curso no site e tragam as dúvidas na próxima aula!!!

<http://www.ic.uff.br/~vanessa>  
(dica: monitorem com <http://www.changedetection.com>)

**Importante:** cadastrem-se em <http://groups.google.com/group/gdse-uff-2011-2>  
(os artigos estão disponíveis lá)

# Datas importantes

---

Data	Atividade
21/09/2011	Apresentação do Trabalho (1a. parte)
23/09/2011	Apresentação do Trabalho (1a. parte)
26/10/2011	PROVA
16/11/2011	Artigo submetido via <a href="#">EasyChair</a> , até 23:55hs
23/11/2011	Apresentação do Trabalho
25/11/2011	Apresentação do Trabalho
	Avaliações dos artigos submetida via <a href="#">EasyChair</a> até 23:55hs
09/12/2011	VS (apenas para graduação)

# Bibliografia

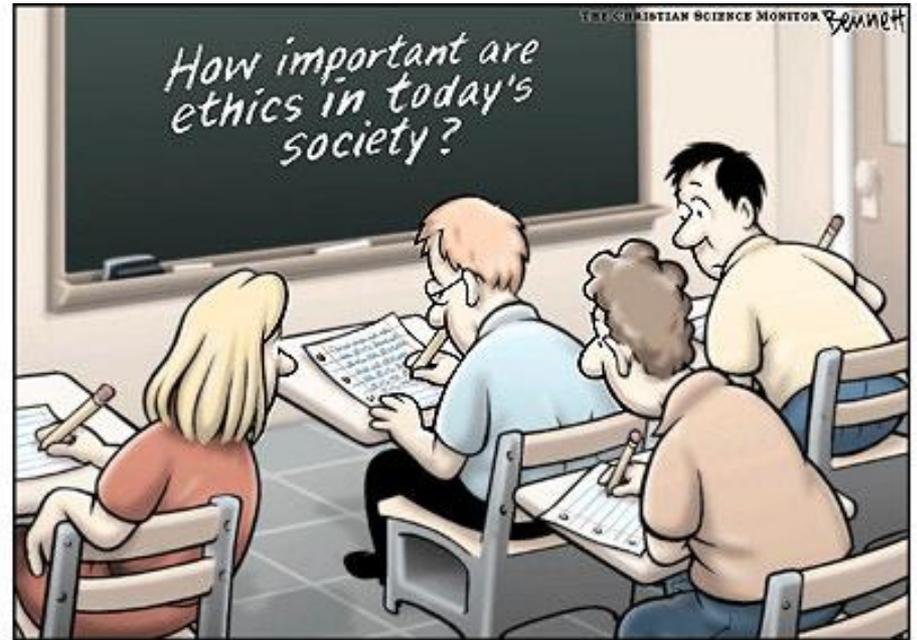
---

- ▶ Abiteboul, Serge e Buneman, Peter. Data On The Web: From Relations To Semistructured Data and XML. Academic Press, 1999.
- ▶ Bradley, Neil. The XML Companion. Addison-Wesley. 3a. Edição, 2001.
- ▶ Chaudhri, Akmal B.; Rashid, Awais e Zicari, Roberto. XML Data Management: Native XML and XML-Enabled Database Systems. Addison-Wesley Professional, 2003.
- ▶ Moro, Mirella e Braganholo, Vanessa. [Desmistificando XML: da pesquisa à prática industrial](#). Atualização em Informática, 2009. Cap. 5. SBC. p. 231-278.
- ▶ Ozu, Nikola; Duckett, Jon; Watt, Andrew , e outros. Professional XML. Peer Information. 2a. Edição, 2001.
- ▶ Especificações de XML: [W3C](#)
- ▶ Tutoriais online: [W3 Schools](#)

# *Fair Play!*

---

- ▶ Não colar ou dar cola em provas
- ▶ Não plagiar o trabalho
- ▶ Não trapacear nas leituras e listas de exercício
- ▶ Não sobrecarregar os colegas do grupo
- ▶ Não assinar presença por colegas
- ▶ Dar crédito apropriado quando usar trabalhos de terceiros



<http://www.claybennett.com/pages/ethics.html>

# Tarefa

---

- ▶ Baixar e instalar o software que iremos usar na disciplina: XML Exchanger Lite (ver link no site da disciplina)