



Gerência de Dados Semiestruturados



Vanessa Braganholo
vanessa@ic.uff.br

Apresentações

- ▶ Um pouco mais sobre mim
 - ▶ Vanessa Braganholo
 - ▶ www.ic.uff.br/~vanessa
- ▶ E vocês?
 - ▶ Nome?
 - ▶ Área de pesquisa? / Orientador?
 - ▶ Período da graduação?
 - ▶ Bolsista? 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 = $(A1 + A2)/2$
- ▶ $A1 = (4 \times \text{Prova} + \text{Avaliação de Artigos}) / 5$
- ▶ $A2 = (4 \times \text{Trabalho} + 2 \times \text{Seminários} + \text{Resumos}) / 7$
- ▶ **APROVADO:** (Presença $\geq 75\%$) E (Média ≥ 6)
- ▶ **VS:** (Aluno de Graduação) E (Presença $\geq 75\%$) E ($4 \leq \text{Média} < 6$)
- ▶ 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!
- ▶ Grupo deve ser o mesmo durante todo o curso
 - ▶ 2 participantes para graduação
 - ▶ Individual para mestrado e 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)
 - ▶ Condução das discussões dos artigos (1 grupo por artigo)
- ▶ Desenvolvimento do trabalho
 - ▶ Apresentação dos trabalhos será dividida em três partes
 - ▶ Proposta
 - ▶ Andamento
 - ▶ Resultado Final

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

Apresentações de Artigo

- ▶ Cada grupo ficará encarregado de apresentar alguns artigos durante o decorrer do curso
 - ▶ Apresentação de 30 minutos para cada artigo
 - ▶ Uso de projetor e/ou quadro branco
- ▶ Os demais grupos ficarão encarregados de defender ou criticar os artigos
- ▶ Um grupo ficará oficialmente como debatedor do artigo, e será o responsável pela condução da discussão

Trabalho

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

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 proposta, andamento e final

- ▶ **Conteúdo do artigo**

- ▶ A motivação e o objetivo do trabalho
- ▶ Trabalhos relacionados (menos ênfase para trabalhos de implementação)
- ▶ Resultados obtidos
- ▶ Considerações finais

Avaliação de Artigos

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

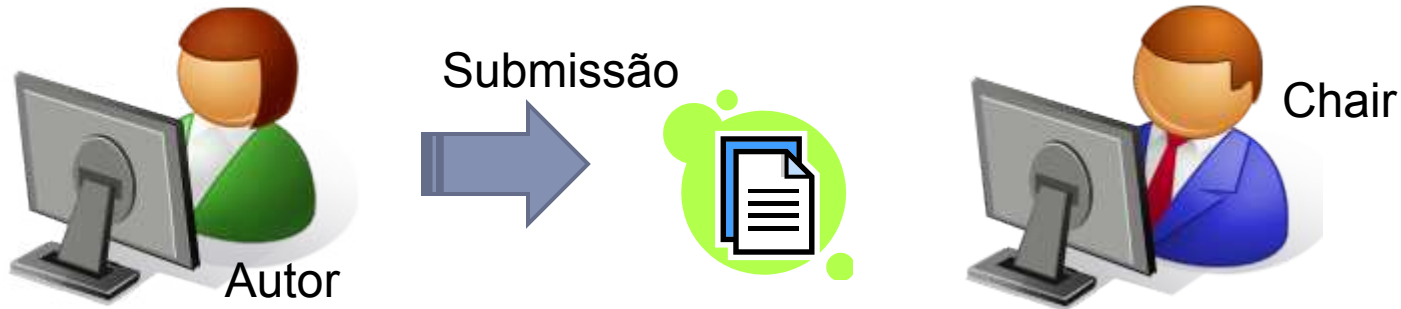
Processo de Avaliação de Artigos



Comitê de Programa
Revisores



Processo de Avaliação de Artigos



Processo de Avaliação de Artigos



Processo de Avaliação de Artigos



Processo de Avaliação de Artigos



Resultado



The word "Resultado" is positioned above a large, light blue arrow pointing from the Chair towards the Author.



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 conta na 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

- ▶ Andamento atual

▶ 3ª 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

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 no grupo do Facebook (link no site da disciplina)

Datas Importantes

- ▶ (vide site da disciplina)

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)