

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Projeto e Especificação de Ontologias

Giovani Rubert Librelotto, PhD
UFSM – Universidade Federal de Santa Maria
librelotto@inf.ufsm.br

1

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria


Sumário Geral

- Introdução a Ontologias
 - Semantic Web
- XML
- Representação do Conhecimento
 - RDF
 - OWL
 - Topic Maps
- Projeto e Construção de Ontologias

2

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria



Parte 2 – XML: Processamento Estruturado de Documentos

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Características do XML

- XML é flexível
 - Meta linguagem
 - Usuários definem suas próprias marcações (*tags*)
- Separação nítida entre conteúdo, estrutura e layout
 - Stylesheets* (XSL) p/ converter em HTML
- Estrutura documentos sob forma de árvore
 - Estrutura pode refletir algum significado
 - APIs p/ *parsers* XML (DOM, SAX)
- Um documento XML pode conter uma descrição opcional de sua gramática (DTD)

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

HTML x XML

HTML:

```
<H1>Introduction to AI</H1>
<UL> <LI>Teacher: Frank van Harmelen
    <LI>Students: 1AI, 1I
    <LI>Requirements: none
</UL>
```

XML:

```
<course>
  <title>Introduction to AI</title>
  <teacher>Frank van Harmelen</teacher>
  <students>1AI, 1I</students>
  <req>none</req>
</course>
```

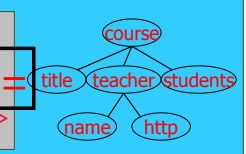
informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Documento XML

- nodo = nome + atributos/valores + conteúdo

```
<course date="...">
  <title>...</title>
  <teacher>...</teacher>
  <name>...</name>
  <http>...</http>
  <students>...</students>
</course>
```



Porém, XML não é suficiente...

- ▶ Vantagem do uso do XML: reutilização do *parser* e validação do documento; porém,
- ▶ Diferentes possibilidades de expressar um mesmo domínio de discurso, podendo acarretar em ambiguidade de interpretação

[illegible]

名詞

εδουατιον

χωρικ

πριωατε

Χς

ναυα

UFSM
Universidade Federal de Santa Maria

informática

Universidade Federal de Santa Maria

Conteúdos

A obra prima

O que toda gente tem ...
e que está sempre a
apontar defeitos ...

O que todos desejam ...
Muitas vezes sem saber bem
como, porquê e o quê.

UFMS
Universidade Federal de Mato Grosso do Sul

informática
Universidade Federal de Mato Grosso do Sul

Conteúdos(2)

A obra prima

O que toda gente tem ...
e aponta defeitos ...

O que todos desejam ...
Muitas vezes sem saber bem
como, porquê e o quê.

informática
Universidade Federal de Santa Maria


Conteúdos

- ▶ Anotação
 - ▶ O que é? Anotar para quê?
- ▶ SGML - a história e o passado
 - ▶ ciclo de vida dos documentos
- ▶ A Evolução forçada pela Web
 - ▶ HTML e SGML: Web SGML
- ▶ XML
 - ▶ Conceitos
 - ▶ Ciclo de vida e Aplicações

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

O que é um documento?




- uma imagem
- uma sequência de palavras
- informação estruturada
- um ser vivo (não no nosso caso)
- ...


UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

O que é que fazemos com eles?



- Eu quero **imprimir** o meu relatório
- O meu chefe quer **pô-lo** na Internet (WWW)
- O meu colega quer **refazer alguns dos cálculos** na folha de cálculo
- Todos querem que fique **seguramente arquivado**




Começa a surgir um problema ...

UFSM
Universidade Federal de Santa Maria


informática
Universidade Federal de Santa Maria

Que "suportes" utilizamos?



- Eu utilizo o **LaTeX**
- O meu colega usa o **Word**
- O nosso chefe ainda usa o **WordPerfect**
- As nossas secretárias utilizam o **Word** no Macintosh
- Ainda há alguém no departamento a usar o **"vi"**

Temos um problema!!!



UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

O Problema:

O que é que necessitamos?

- ☐ Um suporte neutro para a nossa informação
- ☐ Independente de plataformas SW e HW

O que é que o mercado oferecia?

SGML

Standard
Generalized
Markup
Language

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Motivação

- ▶ Maioria da informação em suporte digital: tradicionalmente em MSWord, Wordperfect, Access, Excel, Acrobat.
- ▶ Transferência e manipulação de documentos entre plataformas dispendiosa
- ▶ Manutenção difícil implica uma baixa longevidade
- ▶ Normalização muito difícil
- ▶ Custos de distribuição podem ser altos
- ▶ A migração para a nova tecnologia que se vai apresentar pode não ser uma opção!...

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

SGML: O que é?

- ▶ Uma linguagem para definir linguagens específicas de anotação: SMDL, HTML, DocBook, EAD, TEI.
- ▶ É independente do suporte (HW e SW).
- ▶ Implica a separação do conteúdo e da forma.
- ▶ É um standard ISO
 - ▶ ideal para armazenamento e intercâmbio

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

O que é a Anotação de um texto?

Um meio de tornar explícita uma interpretação desse texto.

Exemplo1: sinais de pontuação
Olá. Tudo bem.
Olá! Tudo bem?

Exemplo2: anotações
"O rei <nome>Afonso Henriques</nome> queria ..."

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Funções da Anotação

- ▶ representar todos os caracteres de um texto
- ▶ identificar a estrutura do texto
- ▶ reduzir o texto a uma ordem linear (árvore)
- ▶ representar informação contextual
- ▶ distinguir o que é texto do que é anotação

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Linguagem de Anotação

- ▶ Especifica como distinguir a anotação do texto
- ▶ Especifica que anotações são necessárias e quais são permitidas
- ▶ Especifica onde as anotações são necessárias e onde são permitidas
- ▶ Define o significado da anotação

O SGML tem estas características todas ... exceto a última

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Documentação Estruturada

- ▶ Vantagens (com anotação descritiva)
 - ▶ portabilidade e intercâmbio entre plataformas
 - ▶ reutilização
 - ▶ armazenamento e consulta
 - ▶ baixos custos de distribuição
 - ▶ manutenção e acesso
 - ▶ longevidade
 - ▶ reconfiguração do aspecto visual

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Documentação Estruturada

- ▶ Desvantagens
 - ▶ grande investimento inicial
 - ▶ capital
 - ▶ tempo
 - ▶ o XML veio aligeirar estes parâmetros...

UFSM
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Anotar um documento

- A tarefa de anotar um documento num sistema de processamento compreende as seguintes fases:
 - analisar a estrutura da informação e os atributos que a caracterizam
 - determinar quais as funções de processamento que produzirão o formato desejado para cada elemento
 - inserir as etiquetas/anotações no texto

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Documento como um todo

1760
Em 14 de Junho deste anno foi mandado sahir immediatamente da Corte o cardeal Acciaiolli e dentro em 4 dias do Reino, para o que se lhe deu decente transporte e acompanhamento. Era Nuncio de sua Santidade o Papa Clemente 13. Aqui principiou a rotura com a Sé Apostolica. Fes-se publico o motivo pela informação que se mandou a Francisco de Almada residente na Corte de Roma, dando-se por fundamento maior o não pôr luminarias por ocasião do casamento que se fes no dia seis de Junho do Senhor Infante D. Pedro com sua sobrinha a Excelentissima Senhora D. Maria, princesa do Brasil, o que o dito Cardeal não fizera, porque lho não fizerão a saber por carta de officio, quando os mais embaixadores sem isso o fizerão. O caso he que querião afasta-lo da Corte e tambem elle não seguia a conduta della por aquelles tempos.

Extrato de "Memórias de Inácio Peixoto dos Santos"

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Documento estruturado

<EXCERTO><TITULO>1760</TITULO>
<PARAG>Em 14 de Junho deste anno foi mandado sahir immediatamente da Corte o cardeal <NOME> Acciaiolli</NOME> e dentro em 4 dias do Reino, para o que se lhe deu decente transporte e acompanhamento. Era Nuncio de sua Santidade o Papa Clemente 13</PARAG>
<PARAG>Aqui principiou a rotura com a <LUGAR>Sé Apostolica</LUGAR>. Fes-se publico o motivo pela informação que se mandou a <NOME> Francisco de Almada</NOME> residente na Corte de Roma, dando-se por fundamento maior o não pôr luminarias por ocasião do casamento que se fes no dia seis de Junho do Senhor <NOME> Infante D. Pedro</NOME> com sua sobrinha a Excelentissima Senhora D. Maria, princesa do Brasil, o que o dito Cardeal não fizera, porque lho não fizerão a saber por carta de officio, quando os mais embaixadores sem isso o fizerão. O caso he que querião afasta-lo da Corte e tambem elle não seguia a conduta della por aquelles tempos.</PARAG></EXCERTO>

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Estrutura do documento

```

graph TD
    raiz --> excerto
    excerto --> titulo
    excerto --> parag
    titulo --> texto1[texto]
    parag --> texto2[texto]
    parag --> nome
    parag --> lugar
  
```

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

O Poema

"Soneto Já Antigo"
(Álvoro de Campos)

Olha, Daisy: quando eu morrer tu há-de dizer aos meus amigos aí de Londres, embora não o sintas, que tu escondes a grande dor da minha morte. Irás de

Londres p'ra Iorque, onde nasceste (dizes que eu nada que tu digas acreditado), contar áquele pobre rapazito que me deu horas tão felizes,

embora não o saibas, que morri... Mesmo ele, a quem eu tanto julguei amar, nada se importará... Depois vai dar

a notícia a essa estranha Cecily que acreditava que eu seria grande... Raios partam a vida e quem lá ande!

(1922)

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Poema: elementos estruturais

"Soneto Já Antigo"
(Álvoro de Campos)

Olha, Daisy: quando eu morrer tu há-de dizer aos meus amigos aí de Londres, embora não o sintas, que tu escondes a grande dor da minha morte. Irás de

Londres p'ra Iorque, onde nasceste (dizes que eu nada que tu digas acreditado), contar áquele pobre rapazito que me deu horas tão felizes,

em bora não o saibas, que morri... Mesmo ele, a quem eu tanto julguei amar, nada se importará... Depois vai dar

a notícia a essa estranha Cecily que acreditava que eu seria grande... Raios partam a vida e quem lá ande! (1922)

Poema: Título, Autor, Corpo, Data
Corpo: Quadra, Quadra, Terno, Terno.
Quadra: Verso, Verso, Verso, Verso
Terno: Verso, Verso, Verso
Verso: (texto | Nome)+
Nome: texto **Título:** texto
Data: texto **Autor:** texto

Os documentos têm estrutura!

► Uma definição contém um termo técnico seguido de um texto descritivo

texto descritivo

termo técnico

definição

novela série de episódios televisivos, inventada no Brasil.

`<!ELEMENT definicao - - (termo, descricao)>`

O texto forma uma hierarquia ordenada

SGML: exemplo

► O elemento do tipo LISTA é formado por um elemento CABEC, seguido por dois elementos do tipo ITEM.

```
<lista>
  <cabec>Os delimitadores podem ser: </cabec>
  <item> explicitos
  <item> inferidos do contexto
</lista>
```

SGML (o início)

► Apesar de todo o poder expressivo e versatilidade não conseguiu se impor

► Grande complexidade

► Grande investimento inicial de capital e de tempo

O HTML é SGML ...

Será que descende de ?

No fim dos anos 80 Tim Berners-Lee misturou:

- um conjunto de *tags* de um DTD utilizado no CERN
- juntou-lhe uma especificação de *estilo*
- e mais um elemento revolucionário: os *links*

O que deu origem ao HTML

HTML (o início)

► A sua grande simplicidade foi o fator determinante para a grande adesão que ele teve;

► A sua grande simplicidade foi também o início de uma época problemática:

- atingiu-se rapidamente o seu limite
- começaram a proliferar extensões não normalizadas

O sentido ecumênico do HTML

- ▶ o HTML representa a linguagem de anotação que as pessoas podem tomar como certa (!!!).
- ▶ quanto mais universal for uma linguagem menos específica ela se torna
 - ▶ quanto maior for o contexto de aplicação menor será a capacidade de expressão semântica.

O cristalizar do HTML

- ▶ o DTD está diretamente implementado nos browsers;
- ▶ a stylesheet também;
- ▶ quem o utiliza não pensa na anotação mas sim na apresentação;

XML (como surgiu?)

O poder do SGML A simplicidade do HTML

XML

Uma linguagem de anotação extensível

XML (conceitos)

Doc. XML

- Bem formado**
 - não há cruzamento de tags
`<A>olá estás ...`
 - pode-se inferir um DTD
 - torna o pós-processamento mais específico.
- Válido**
 - pertence a uma classe (DTD ou XML Schema)

SGML → XML

- ▶ Fechar todas as tags que forem abertas
- ▶ Converter os elementos vazios
 - ▶ `<HR> ==> <HR/>`
- ▶ Colocar os valores dos atributos dentro de aspas
 - ▶ ``

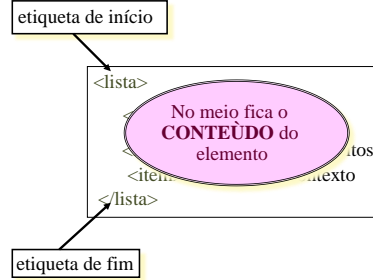
Um documento XML compreende

- ▶ uma declaração XML
 - ▶ `<?xml version="1.0" ... ?>`
- ▶ um esquema em DTD ou XML Schema (opcional)
- ▶ a instância do documento

XML: exemplo - instância

```
<?xml version="1.0"?>
<!DOCTYPE poema SYSTEM "poema.dtd">
<poema>
  <titulo>"Soneto J&#225; Antigo"</titulo>
  <autor>(&#193;lvaro de Campos)</autor>
  <corpo>
    <quadra>
      <verso>Olha, <nome>Daisy</nome>; quando eu morrer tu h&#225;s de</verso>
      <verso>dizer aos meus amigos a&#237; de <nome>Londres</nome>;</verso>
      <verso>embora n&#227;o o sintas, que tu escondes</verso>
      <verso>a grande dor da minha morte. In&#225;s de</verso>
    </quadra>
    <quadra>
      <verso><nome>Londres</nome> pra <nome>lorque</nome>, onde nasceste (dizes</verso>
      <verso>que eu nada que tu digas acredito);</verso>
      <verso>contar &#224;quele pobre rapazito</verso>
      <verso>que me deu horas t&#227;o felizes;</verso>
    </quadra>
  </corpo>
</poema>
```

Ocorrência de um elemento



DTD (Definição do Tipo de Documento)

- Os tipos de elementos permitidos dentro do documento
- As características de cada tipo de elemento, inclusive os atributos permitidos e o conteúdo que ele pode ter
- As notações e entidades que podem ser encontradas dentro do documento

XML: exemplo - DTD do poema

```
<!-- declarações para os elementos -->
<!ELEMENT poema (titulo, autor, corpo, data) >
<!ELEMENT titulo (#PCDATA)>
<!ELEMENT autor (#PCDATA)>
<!ELEMENT corpo ((quadra|terno)+)>
<!ELEMENT quadra (verso, verso, verso, verso)>
<!ELEMENT terno (verso, verso, verso)>
<!ELEMENT verso (#PCDATA|nome) +>
<!ELEMENT nome (#PCDATA) >
<!ELEMENT data (#PCDATA) >
```

DTD: definição de um elemento

```
<!ELEMENT lista (cabec, item+) >
<!ELEMENT item (#PCDATA) >
```

DTD: Definição do Conteúdo

- outros elementos especificados
- ANY (qualquer elemento especificado)
- EMPTY (nada, vazio)
- #PCDATA (texto)
- uma mistura de elementos com #PCDATA (conteúdo misto)

Expressão de Conteúdo: sintaxe

- ▶ sequência
 - ▶ a,b a seguido de b
 - ▶ a|b a ou b, mas não ambos
 - ▶ a&b equivalente a (a,b) | (b,a)
- ▶ ocorrência
 - ▶ a um e apenas um
 - ▶ a? opcionalmente um (0 ou 1)
 - ▶ a* zero ou mais
 - ▶ a+ um ou mais

Exemplo: o poema

Poema: título, autor, corpo, data
 Corpo: quadra, quadra, terno, terno.
 Quadra: verso, verso, verso, verso
 Terno: verso, verso, verso
 Verso: (texto | nome)+
 Nome: texto

```
<!ELEMENT poema (título,autor,corpo,data) >
<!ELEMENT corpo (quadra,quadra,terno,terno) >
<!ELEMENT quadra (verso,verso,verso,verso) >
<!ELEMENT terno (verso,verso,verso) >
<!ELEMENT verso (#PCDATA |nome)* >
```

Um elemento pode ter atributos

- ▶ para conter informação para além do tipo e do contexto
- ▶ para identificação de ocorrências específicas de elementos
- ▶ para fazer algumas validações (poucas)

nome do atributo valor do atributo

```
<lista tipo="bola" id="L123">
  <item id="L123.1">delimitadores explicitos</item>
  <item id="L123.2">inferidos do contexto</item>
</lista>
```

Ex: anotação morfo-sintáctica

```
<quadra>
<verso><verbo tempo="imperativo"
  pessoa="2s">Olha</verbo>,<nome> Daisy</nome>; quando eu morrer tu
  hás-de</verso>
<verso><verbo tempo="infinitivo">dizer</verbo> aos meus amigos aí
  de <nome>Londres</nome>,</verso>
<verso>embora não o sintas, que tu escondes</verso>
<verso>a grande dor da minha morte. Irás de</verso>
</quadra>
```

Definição de um atributo

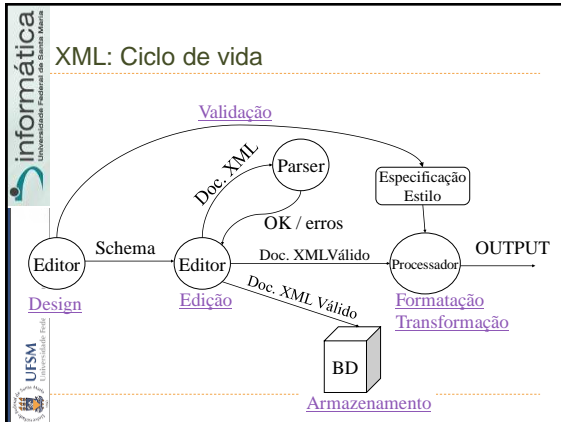
- ▶ os identificadores de nomes e tipos devem ser únicos dentro de um elemento

nome tipo valor por omissão

```
<!ATTLIST lista
  id ID #IMPLIED
  tipo (simplea|bola) bola >
```

Valores possíveis

- ▶ #REQUIRED (obrigatório)
- ▶ #FIXED (constante)
- ▶ #IMPLIED (opcional)
- ▶ valor explícito



XML Schema

- ▶ **Motivação**
 - ▶ Sintaxe XML
 - ▶ Suporte para NameSpaces
 - ▶ Criação de tipos de dados
 - ▶ Poder referenciar tipos de dados já criados
 - ▶ Permitir estender outros tipos
 - ▶ Restringir tipos de dados existentes
 - ▶ Permitir o uso de ferramentas gráficas
 - ▶ ...

Tipos Primitivos

- ▶ String
- ▶ Float
- ▶ Decimal
- ▶ Date
- ▶ ...

```

<xs:element name="nome" type="xs:string"/>
<xs:element name="lugar" type="xs:string"/>
<xs:element name="idade" type="xs:decimal"/>
<xs:element name="data" type="xs:date"/>
  
```

Tipos Simples e Complexos

- ▶ Os tipos de dados em XML Schema são "simple" ou "complex".
- ▶ Um tipo "simple" é um dos tipos básicos: string, date, float, double, timeDuration, ...
- ▶ Um elemento que tenha atributos ou elementos filho é do tipo "complex".

Exemplo: complex Type

```

<xs:complexType name="Tquadra">
  <xs:sequence>
    <xs:element name="verso" type="Tverso"/>
    <xs:element name="verso" type="Tverso"/>
    <xs:element name="verso" type="Tverso"/>
    <xs:element name="verso" type="Tverso"/>
  </xs:sequence>
</xs:complexType>
  
```

Tquadra é um *complex type*

Tverso é uma referência a um *complex type* definido em algum lugar no esquema

Scope dos elementos

- ▶ Um elemento pode ser declarado como global ou como local:
 - ▶ Um elemento global é declarado como sendo filho de <schema>.
 - ▶ Os elementos globais podem ser reutilizados por referência.
- ▶ Um elemento local é declarado em algum lugar na estrutura do Schema.

informática
Universidade Federal de Santa Maria

Cardinalidade

```
<xs:complexType name="Tquadra">
  <xs:sequence>
    <xs:element name="verso" type="Tverso"
      minOccurs="4" maxOccurs="4"/>
  </xs:sequence>
</xs:complexType>
```

A cardinalidade pode ser restringida recorrendo aos atributos: **minOccurs** e **maxOccurs**

UFSC
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Valores para elementos

```
<xs:element name="email" type="xs:string"
  default="giovani@unifra.br"/>
```

Valor por default

```
<xs:element name="supervisor" type="xs:string"
  fixed="Giovani Rubert Librelotto"/>
```

Valor fixo

UFSC
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Atributos

- Os atributos declaram-se recorrendo ao elemento `<attribute>`
- Um atributo é sempre do tipo simple

UFSC
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Restrição de tipos primitivos

```
<xs:simpleType name="dia" base="xs:integer">
  <xs:minInclusive value="1"/>
  <xs:maxInclusive value="31"/>
</xs:simpleType>
```

Restrição de domínio

```
<xs:simpleType name="telefone" base="xs:integer">
  <xs:pattern value="332\d(1).\d{4}"/>
</xs:simpleType>
```

Restrição de formato

UFSC
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

Restrição por enumeração

```
<xs:simpleType name="diaSemana" base="xs:string">
  <xs:enumeration value="Domingo"/>
  <xs:enumeration value="Segunda"/>
  <xs:enumeration value="Terça"/>
  <xs:enumeration value="Quarta"/>
  <xs:enumeration value="Quinta"/>
  <xs:enumeration value="Sexta"/>
  <xs:enumeration value="Sábado"/>
</xs:simpleType>
```

UFSC
Universidade Federal de Santa Maria

informática
Universidade Federal de Santa Maria

A estrutura do nosso poema

```

graph TD
    Poema --> soneto
    soneto --> titulo
    soneto --> autor
    soneto --> corpo
    soneto --> data
    corpo --> quadra1[quadra]
    corpo --> quadra2[quadra]
    corpo --> temo1[temo]
    corpo --> temo2[temo]
    quadra1 --> verso1[verso]
    quadra1 --> verso2[verso]
    quadra1 --> verso3[verso]
    quadra1 --> verso4[verso]
    quadra2 --> verso5[verso]
    quadra2 --> verso6[verso]
    quadra2 --> verso7[verso]
    quadra2 --> verso8[verso]
    temo1 --> verso9[verso]
    temo1 --> verso10[verso]
    temo1 --> verso11[verso]
    temo1 --> verso12[verso]
    temo2 --> verso13[verso]
    temo2 --> verso14[verso]
    temo2 --> verso15[verso]
    temo2 --> verso16[verso]
  
```

UFSC
Universidade Federal de Santa Maria

Extensible Stylesheet Language (XSL)

- Porque Stylesheets?
 - Separação do conteúdo (XML) da apresentação (XSL)
- Porque não usar apenas CSS para XML?
 - XSL é muito mais poderoso:
 - seleção de elementos
 - transforma a árvore XML
 - conteúdo baseado (resultado pode depender dos valores atuais)

XSL(T) Overview

- Folhas de estilo XSL são denotados em sintaxe XML
- Componentes XSL:
 1. uma linguagem para transformar documentos XML (XSLT: parte integral da especificação XSL)
 2. Um vocabulário de formatação XML (Formatting Objects (FO): >90% das propriedades de formatação herdada de CSS)

Modelo de Processamento XSLT

Árvore inicial XML

Transformation
XSLT stylesheet

XML, HTML, csv, text...
árvore gerada

XSL e normas associadas

XSL

XSLFO

XSLT

XPath

Cabeçalho XSL

Método de saída

~ main() em C

Título do cabeçalho HTML

Instruções HTML

Título do poema em H1

Para cada verso de cada quadra do corpo do poema, imprimir o conteúdo do verso.

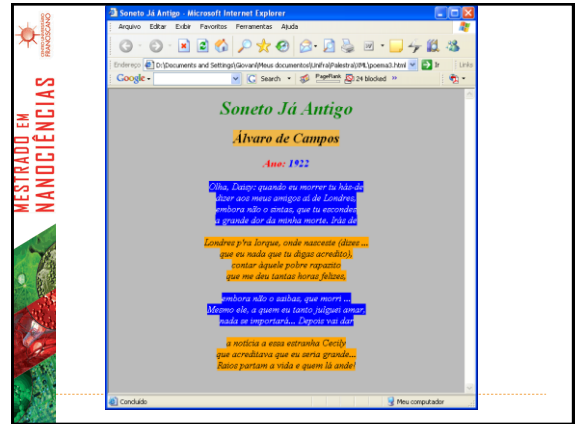
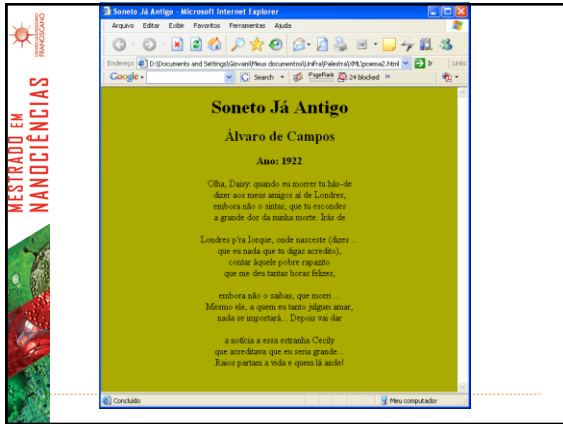
Instrução XPath

Soneto Já Antigo

Álvaro de Campos

Ano: 1922

Olha, Diário quando eu morrer tu há-de
dizer aos meus amigos aí de Londres,
embora não o sintas, que tu recordas
a grande dor da minha morte. Já de
Londres p'va Iorque, onde nascente (dizes...
que eu nada que tu digas acrédio),
contar alguns poetas rapado
que me deu tantas horas felizes,
embora não o saibas, que morri
Mesmo ele, a quem eu tanto julgar amar,
nada se importará. Depois vai dar
a notícia a essa estranha Cecily
que acreditava que eu seria grande.
E não partas a vida e quem lá sabe!



Mecanismos XSL

- ▶ Todos os recursos HTML
 - ▶ Links
 - ▶ Tabelas
 - ▶ Imagens
 - ▶ ...
- ▶ Criação de índices
- ▶ XSL: 1 fonte para N objetos
- ▶ Etc...

XPath

- ▶ A sintaxe básica do XPath é muito semelhante à do endereçamento de arquivos em um sistema operacional.
- ▶ Se o endereço começar por /, então estaremos perante um endereço absoluto, ou seja, a raiz.

XPath em XSL

- ▶ Em termos funcionais, a utilização do XPath em XSL pode ter uma série de objetivos:
 - ▶ Seleção de nodos para processamento;
 - ▶ Especificação de condições permitindo diferentes modos de processamento de um mesmo nodo;
 - ▶ Geração de texto a ser incluído na árvore final.

Nódo raiz: `/poema`

Nódo atributo: `/poema/@id`

Nódo elemento: `/poema/autor`

Todos os nomes: `//nome`

3º verso da 2ª quadra: `//quadra[2]/verso[3]`

Lugares que começam com I: `starts-with(/lugar, 'I')`

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Funções XSL

- ▶ Funções que permitem manipular texto.
 - ▶ concat, starts-with, contains, substring, ...
- ▶ Funções que produzem um resultado booleano.
 - ▶ boolean, not, true, false, ...
- ▶ Funções que nos permitem realizar cálculos aritméticos com o conteúdo dos elementos.
 - ▶ number, sum, floor, ceiling, ...

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Linguagens XML

- ▶ Linguagens definidas em SGML/XML
 - ▶ DocBook XML 1.5
 - ▶ SMDL - música
 - ▶ CML - química
 - ▶ SMIL - multimídia e interface
 - ▶ MathML - matemática
 - ▶ H7 - medicina
 - ▶ XML-EDI - comércio eletrônico

informática
Universidade Federal de Santa Maria

UFSM
Universidade Federal de Santa Maria

Referências...

- ▶ <http://www.w3.org/XML/>
- ▶ <http://www.xml.com>
- ▶ <http://www.jclark.com/>
- ▶ <http://www.oasis-open.org/cover/>
- ▶ <http://www.di.uminho.pt/~jcr/>
- ▶ ...