

# Guião 3: XML

Informática Aplicada – Linguagens de Anotação de Documentos

Abril 2017

## Objetivos

Explorar linguagens de anotação para estruturar dados. Explorar a representação e manipulação de dados em XML.

## Descrição

A representação de informação de modo estruturado é essencial para a partilha de dados de forma precisa pelas várias ferramentas. Para que possam ser processados de modo uniforme, mas ainda legíveis por humanos, linguagens de anotação são frequentemente usadas para serializar e armazenar os dados.

O XML é uma linguagem de anotação para a representação de informação arbitrária numa estrutura em árvore, não tendo qualquer esquema predefinido. O XML permite ao utilizador definir as suas próprias anotações, específicas para o seu domínio de aplicação, aplicando apenas regras restritas sobre a forma dessas anotações. Isto torna o XML extremamente flexível mas ainda facilmente processável.

Por vezes é necessário acordar o esquema dos documentos XML entre as partes interessadas, para garantir a representação e interpretação uniforme dos dados. Em XML, DTDs podem ser atribuídos aos documentos que definem o que deve ser considerado um documento válido, através da declaração dos elementos e atributos válidos, assim como o seu conteúdo.

Uma vantagem de representar dados de modo estruturado é a possibilidade de fazer consultas automáticas sobre a informação, tarefas que seriam trabalhosas manualmente. Para documentos XML, o XPath é uma linguagem de consulta básica que permite a consulta de dados, através da definição de expressões de navegação da árvore do documento.

Dados estruturados são também mais facilmente transformados para diferentes outros tipos de dados para que possam ser utilizados em diferentes contextos de aplicação. No contexto do XML, isto pode ser feito usando a linguagem XSLT, que define transformações sobre elementos de um documento XML.

## Tarefas

As seguintes tarefas deverão ser elaboradas no desenvolvimento deste trabalho:

**Estrutura de dados** Considere os dados recolhidos pelo formulário desenvolvido durante o Guião 1, persistidos numa folha de cálculo em formato tabular. Defina um modelo conceptual em árvore para representar esses dados.

**DTD** Crie um esquema DTD que defina a estrutura de dados em árvore conceptualizada no ponto anterior.

**XML** Crie um documento XML, válido segundo o DTD definido no ponto anterior, que represente os dados recolhidos pelo formulário.

**XPath** Defina algumas expressões XPath de consulta básicas sobre a estrutura de dados definida (e.g., todas as respostas com uma certa percentagem de escolha).

**XSLT** Defina uma transformação XSLT básica que converta os dados XML numa página HTML.

**Relatório** Escreva um pequeno relatório em  $\text{\LaTeX}$  que descreva os artefactos definidos nos pontos anteriores<sup>1</sup>. Adicione o ficheiro XML, com o DTD definido internamente, assim como o ficheiro XSLT, à diretoria do grupo no Google Drive.

## Submissão

O trabalho deve ser desenvolvido em grupos de 3, e a versão dos documentos que será avaliada é a que estiver presente no *Google Drive* às 9h00 do dia 24 de Abril. Não serão aceites submissões por outro meio nem após a data definida.

---

<sup>1</sup>Pode usar o ambiente `verbatim` para excertos de código.