**Capa** - identificação do tema trabalho, a disciplina em que se enquadra, curso, ano lectivo,

turma prática laboratorial e a identificação completa dos alunos do grupo.

**Índice**

**Introdução** – Identificação e descrição do trabalho e objectivos propostos.

**Análise** – Especificação da análise do domínio do problema proposto.

3

**A Linguagem** - Definição da sintaxe da linguagem proposta através de um XMLSchema e

criação de um documento XML instância válida dessa linguagem.

**Transformações** – Descrição de StyleSheets (XSLT) adequadas à transformação da

informação do documento XML em novos formatos.

**Conclusão**

**Anexos**

**Introdução**

O modo de guardar informação nos tempos que correm ainda é uma temática complexa. Há vários formatos concorrentes, como por exemplo, JSON (JavaScript Object Notation) ou XML (eXtensible Markup Language), mas para o este trabalho, foi-nos imposto a obrigatoriedade da utilização do XML.

O XML, sendo base para outros formatos de ficheiro, como o [XHTML](http://pt.wikipedia.org/wiki/XHTML) (eXtensible Hypertext Markup Language), foi-se tornando uma linguagem altamente poderosa.

Assim, foi-nos proposto a criação de um XML corretamente validado por um XSD (XML Schema Definition), e que seja capaz de fazer a gestão de um Jornal Desportivo, passível de ser transformado noutros formatos de ficheiro, nomeadamente, HTML (HyperText Markup Language) e também com algumas filtragens (por classificação na tabela, por clube, etc.).

**Analise**

A necessidade de nos sentirmos informados, fez com que tivesse que se encontrar uma alternativa viável. Para isso, criou-se o jornal generalista em que todas as “novidades” seriam impressas num papel. Depois disso, houve quem decidisse não fazer um jornal generalista, mas sim jornais por tema (desportivos, financeiros), de maneira a ter mais noticias relacionadas com aquele tema e para atrair público-alvo. Com a evolução da tecnologia, o papel começa a entrar em desuso e cada vez mais, os adeptos dos jornais usam o seu tablet ou computador para aceder ao site de notícias favorito.

Nesse contexto surgiu a necessidade de haver uma organização de informação, de maneira a proporcionar ferramentas melhoradas, quer de filtragens, quer de pesquisas.

Cada notícia desportiva, é escrita numa certa língua, podendo ou não ser traduzida em várias línguas (geralmente acontece apenas nos jornais internacionais), também terá o seu autor de notícia, que geralmente é um jornalista associado ao jornal, embora haja também quem redija opiniões públicas e as mesmas aparecem no jornal. No caso da data, optou-se por colocar a data de emissão da notícia e não a data de escrita da notícia.

Relativamente ao conteúdo da notícia, pode-se dizer que cada noticia tem o seu clube, que por sua vez tem uma classificação associada no campeonato em que participa que é de um determinado desporto, tendo este clube uma sede num determinado pais. Ainda relativamente ao clube, pode-se dizer que geralmente um clube tem uma sigla associada, por exemplo, quando dizemos SLB, claramente nos referimos ao Sport Lisboa e Benfica, sediado em Portugal.

Para facilitar a pesquisa e para ser mais fácil relacionar uma notícia com a outra, optou-se por atribuir um identificador à noticia.

**Linguagem**

Para este problema, cada notícia possui um atributo id (com o formato AAAAA-9999-9999, sendo os AAAAA, qualquer letra de A-Z, com tamanho variável de 2 a 5, e os 9999, qualquer algarismo de 0-9, com tamanho fixo de 4, sendo que o primeiro conjunto de 9999 se opte por colocar o ano de publicação da notícia). Esse mesmo id será também utilizado para referenciar as notícias relacionadas e similares com a notícia principal.

Cada noticia terá os seguintes elementos: data/tempo, autor, fonte, titulo, abstract, fotos, clube, desporto, campeonato, outras notícias e corpo da notícia, que irão ser explicados de seguida.

A data/tempo é dividido em 5 elementos do tipo inteiro, sendo eles: ano, mês, dia, hora e minutos. Pensamos que não houve a necessidade de inserção dos segundos, dado que por vezes isso é omitido nas principais fontes de informação.

O autor, possui 2 elementos, sendo eles o nome e email. O nome é do tipo texto, enquanto que o email contém uma validação para verificar se o email introduzido é o correto.

A fonte contem 2 elementos sendo eles o nome da fonte e o URI da fonte, sendo o nome do tipo texto e o URI do tipo URI.

O título, abstract e corpo da notícia, são todos do mesmo tipo. Como foi-nos pedido que a notícia poderia estar em várias línguas, estes elementos possuem subelementos que especificam em que língua é que estão escritos, sendo apenas possível definir noticias nas línguas portuguesa, inglesa, francesa, italiana e espanhola. Para se adicionar uma nova língua, terá que se adicionar ao tipo de línguas, o elemento respetivo.

As fotos contém dois elementos diversificados a imagem principal e outras imagens, mas com os mesmos atributos, sendo eles o href (irá informar a localização da imagem), width e height (tamanho respetivo da imagem). A diferença prende-se com o número de imagens a poderem ser colocadas, sendo que só poderá haver uma imagem principal e um número ilimitado de outras imagens.

O clube contem 4 elementos, sendo eles o nome, o nome pequeno, o país e a posição atual no campeonato. O elemento nome é do tipo texto e o elemento posição é do tipo inteiro. O nome pequeno tem uma limitação de que só aceita entre 3 a 5 letras maiúsculas enquanto que o pais só aceita 3 letras maiúsculas.

O elemento desporto é do tipo texto, não havendo nenhuma restrição.

O elemento campeonato tem um atributo localização, em que deverá ser colocado o código do país em 3 letras maiúsculas, e o elemento nome que é do tipo string.

O elemento outras notícias, subdivide-se em 2 elementos iguais, o noticias relacionadas e similares. Em ambos os casos, eles aceitam o elemento relacionada ou similar (conforme o elemento que estivermos a falar) com um atributo, sendo ele o id da notícia a que nós nos estamos a referir, estando em conformidade com o id da notícia.

**Transformações**

**Conclusão**