O que é um xml? É uma linguagem de marcação que tem o objetivo de ser lida por computadores e pessoas. Ela utiliza nodos como se fosse uma arvore onde cada marcador pode ter um filho e ou ter um pai. Assim o acesso é dado como uma forma indireta passando pelo pai até o filho procurado. Temos uma raiz também como qualquer arvore estrutural.

Primeiro problema encontrado: foi a acentuação dos xml, abrindo ele com browser é possível encontrar uma palavra que costuma ter acentuação estava impressa com lixo ao invés de imprimir o carácteres acentuado. Em algumas pesquisas na web foi possível ver que esse problema tem origem na encouding do xml. Que estava em "ISO-8859-1" esse formato não aceita carácteres acentuados, a solução para o trabalho foi necessário trocar a encouding dos xml para “UTF-8”. Assim o problema foi solucionado. Como foi feita a troca? Ao invés de criarmos uma função que corrija as encoudings. Devido a definição do trabalho era baixar os currículos em xml. Era mais rápido trocar a encouding um por um do que pensar em um algoritmo para a trocar todas as encoudings sempre quando for lido o xml.