

Základní pravidla struktury dokumentů XML

- každý XML dokument musí obsahovat alespoň jeden element
- XML dokumenty musí začínat XML deklarací, která určuje, jaké verze XML a jakého typu kódování znaků bylo použito; implicitním kódováním dokumentů je Unicode (UTF-8), je však možno určit v XML deklaraci jakékoliv jiné kódování (pro češtinu ISO-2 nebo WIN-1250)
- dokument musí obsahovat pouze jeden element nejvyšší úrovně – tzv. kořenový element (všechny ostatní elementy musí být obsaženy v něm)
- elementy musí být správně zanořeny (pokud element začíná uvnitř jiného elementu, musí být v témž elementu také ukončen), nesmí dojít k tzv. křížení elementů
- každý element musí obsahovat jak počáteční, tak ukončovací značku (tag); jedinou výjimkou je tzv. prázdný element (element bez obsahu), v rámci struktury stojí někde mezi elementem a atributem
- název elementu v počáteční značce musí být totožný s názvem ve značce ukončovací; v názvech elementů jsou rozlišována velká a malá písmena
- název elementu může obsahovat písmena, číslice, podtržítka, oddělovník, dvojtečku nebo tečku; dvojtečka může být použita pouze ve speciálním případě, kdy odděluje tzv. jmenný prostor; název elementu může začínat pouze písmenem, podtržítkem nebo dvojtečkou
- komentáře se mohou vyskytovat na libovolném místě v dokumentu mimo vlastní značkování; řetězec "--" se nesmí objevit uvnitř komentáře
- každý element může mít žádný, jeden nebo několik atributů; pro jejich názvy platí stejná pravidla jako pro názvy elementů; jméno atributu je odděleno od jeho hodnoty rovnítkem "="; hodnota atributu musí být uvedena uvnitř apostrofů '.' nebo uvozovek "..."