

XML (Extensible Markup Language)

XML je rozšiřitelný značkovací jazyk, který definuje sadu pravidel pro kódování dokumentů ve formátu čitelném pro lidi i stroje. Na rozdíl od HTML, které je zaměřeno na **zobrazení** dat, XML je navrženo výhradně pro **přenos a ukládání** dat.

XML je standardem organizace W3C a tvoří základ mnoha dalších technologií, jako jsou SVG, RSS, SOAP nebo kancelářské formáty (DOCX, XLSX).

Základní struktura XML

Každý XML dokument musí být tzv. „dobře strukturovaný“ (well-formed). To znamená, že musí splňovat následující pravidla:

- **Deklarace:** Dokument by měl začínat řádkem definujícím verzi a kódování.
- **Kořenový element:** Musí existovat právě jeden prvek, který obaluje celý zbytek dokumentu.
- **Párové značky:** Každý element musí mít počáteční a koncovou značku (např. <tag>...</tag>).
- **Správné vnořování:** Značky se nesmí překrývat.
- **Case sensitivity:** XML rozlišuje malá a velká písmena (<Uživatel> není totéž co <uživatel>).

Ukázka XML dokumentu

```
<?xml version="1.0" encoding="UTF-8"?>
<knihovna>
  <kniha id="101">
    <nazev>Design Patterns</nazev>
    <autor>Erich Gamma</autor>
    <rok>1994</rok>
  </kniha>
  <kniha id="102">
    <nazev>Clean Code</nazev>
    <autor>Robert C. Martin</autor>
    <rok>2008</rok>
  </kniha>
</knihovna>
```

Klíčové komponenty

1. Elementy (Prvky)

Hlavní datové jednotky ohraničené značkami. Mohou obsahovat text, další elementy nebo obojí.

2. Atributy

Doplňující informace o elementu zapsané uvnitř počáteční značky (např. `id=„101“`). Atributy by měly obsahovat metadata, zatímco samotná data patří do elementů.

3. Jmenné prostory (Namespaces)

Zabraňují konfliktům jmen, pokud se v jednom dokumentu kombinují data z různých zdrojů. Definují se pomocí atributu `xmlns`.

Validace: DTD a XML Schema (XSD)

Aby bylo zajištěno, že XML dokument obsahuje správná data ve správném formátu, používají se schémata:

- **DTD (Document Type Definition):** Starší a jednodušší způsob definice struktury.
- **XSD (XML Schema Definition):** Moderní způsob, který sám používá XML syntaxi. Umožňuje definovat datové typy (číslo, datum, řetězec), omezení délek nebo regulární výrazy.

Související technologie

Technologie	Popis
XPath	Jazyk pro navigaci a výběr částí XML dokumentu (např. <code>/knihovna/kniha[1]</code>).
XSLT	Jazyk pro transformaci XML do jiných formátů (HTML, PDF, jiný XML).
DOM	Model, který načte celé XML do paměti jako stromovou strukturu (náročné na RAM).
SAX	Událostmi řízený parser, který čte XML postupně (vhodné pro obří soubory).

Srovnání: XML vs JSON

Vlastnost	XML	JSON
Lidská čitelnost	Dobrá (upovídaná)	Velmi dobrá (stručná)
Velikost souboru	Větší (kvůli uzavíracím značkám)	Menší
Datové typy	Vše je text (nutno definovat v XSD)	Podporuje čísla, pole, objekty, boolean
Podpora polí	Musí se simulovat opakováním elementů	Nativní podpora []

Zajímavost: Formáty Microsoft Office (např. `.docx`) jsou ve skutečnosti přejmenované archivy ZIP, které uvnitř obsahují velké množství propojených XML souborů definujících text, formátování a obrázky.

[Zpět na Data](#)

From:
<http://serviceit.cz/> - **IT ENCYKLOPEDIE**

Permanent link:
<http://serviceit.cz/doku.php?id=xml>

Last update: **2025/12/31 14:22**

