

Kódovanie textu

Spôsob kódovania sa volá **ASCII** – *American Standard Code for Information Interchange* (Americký štandardný kód pre výmenu informácií).

Tento štandard hovorí, že na zakódovanie každého znaku sa použije 7 bitov. Čo umožňovalo definovať kód pre 128 znakov. Neskôr sa kódovanie rozšírilo na 8 bitov a 256 znakov. Pritom prvá polovica znakov je pre všetky krajiny rovnaká a zvyšných 128 znakov sa pre každú krajinu stanovil podľa ich potrieb. Tento spôsob vniesol do kódovania znakov chaos (veľmi veľké množstvo regionálnych znakových sád), navyše pre znaky niektorých písom 7-bitový priestor nestačil. Vznikol preto nový spôsob kódovania **UNICODE**.

Toto kódovanie používa 16 a 24 bitov na zakódovanie jedného znaku, čo umožňuje zakódovať cca 1,1 milióna možných znakov. Tento počet znakov umožňuje zakódovať znaky všetkých relevantných abecied pomocou jednej medzinárodnej tabuľky. Toto kódovanie zabezpečuje, že ten istý znak má rovnaký kód v každej krajine i na každom type počítača.

Nevýhodou tohto kódovania je, že znaky, ktoré sme predtým vedeli zakódovať iba ôsmimi bitmi (jedným bajtom) v ASCII, v Unicode sú kódované 16 a 24 bitmi (dvomi a tromi bajtami), a teda zaberajú viac pamäte ako by zaberá text v kódovaní ASCII. Istým riešením tohoto problému je formát **UTF-8**. V tomto kódovaní je prvých 128 znakov tabuľky ASCII (tieto sú pre všetky krajiny rovnaké) zakódovaných pomocou 8 bitov a zvyšné znaky sú zakódované 16 a 24 či 32 bitmi. Toto kódovanie je výhodné pre anglicky hovoriace krajiny a krajiny, v ktorých väčšinu znakov textu tvorí len prvých 128 znakov tabuľky ASCII.

Neformátovaný a formátovaný text

Formátovaný text je text uložený do súboru vo forme, kde sú uložené aj informácie o formátovaní (fonty, veľkosť písma, nadpisy, odkazy...). Doplnujúce informácie sú uložené buď v čitateľnej podobe (výsledný súbor je textový – napríklad HTML, XML, ktorý používajú open office, TeX, microsoft office) alebo sú formátovacie informácie uložené v nečitateľnej podobe ako binárne dáta, ktoré neodpovedajú tlačným znakom (súbory .doc, .xls a ďalšie z Microsoft Office).

Opakom formátovaného textu je **jednoduchý text**, ktorý formátovacie informácie neobsahuje.

Typy súborov s textom:

RTF, PDF – adobe acrobat reader, DOC - word, DOCX - word, ODT - writer, TXT – textový súbor (poznámkový blok)...

Značkovací jazyk:

HTML – jazyk na tvorbu www stránok, XML – jazyk pre použitie textového súboru v rôznych textových editoroch, TeX – profesionálny textový editor

Prevod do PDF a z PDF

Textový editor má v ponuke možnosť uloženia dokumentu vo formáte PDF. Treba hľadať v ponuke „uložiť ako“. Na opačný prevod (PDF do doc...) môžeme použiť rôzne online prevodníky, prípadne freeware alebo profesionálne platené prekladače.

Kancelárske balíky

Microsoft Office, Open Office, Libre Office, Google docs, T602...