

UTF (Unicode Transformation Format)

UTF je rodina standardů pro kódování znaků, která umožňuje reprezentovat jakýkoliv znak z mezinárodní sady **Unicode** (včetně české diakritiky, azbuky, čínských znaků nebo emoji) v digitální podobě. Díky UTF můžeme v našich systémech bezpečně pracovat s textem v jakémkoliv jazyce bez rizika vzniku nečitelných znaků (tzv. „rozsypaný čaj“).

Formáty UTF používané v naší síti

1. UTF-8

Nejpoužívanější formát v naší společnosti a na celém internetu.

- **Variabilní délka:** Jeden znak může zabírat 1 až 4 bajty.
- **Zpětná kompatibilita:** Znaky základní anglické abecedy (ASCII) zabírají pouze 1 bajt, což šetří místo.
- **Využití:** Standard pro naše **WWW** servery, konfigurační soubory na **VPS** a zdrojové kódy v rámci **VCS**.

2. UTF-16

Znaky jsou kódovány primárně pomocí 2 nebo 4 bajtů.

- **Využití:** Interní reprezentace textu v operačních systémech Windows a v programovacích jazycích jako Java nebo JavaScript, které používá náš **Vývojový tým**.

Proč je UTF klíčové pro naše systémy?

Správné nastavení UTF kódování zajišťuje:

- **Integritu dat:** Zabraňuje poškození jmen zákazníků s diakritikou v našich databázích.
- **Konzistenci v Jira:** Umožňuje kolegům z mezinárodních poboček sdílet informace bez ohledu na jejich lokální klávesnici.
- **Lokalizaci:** Naše **Marketingové oddělení** může připravovat materiály pro globální trhy pomocí jednotného standardu.

Správa a konfigurace (Best Practices)

Aby naše **IT Podpora** předešla problémům s kódováním, dodržujeme tato pravidla:

1. ****"UTF-8 Everywhere":**** Všechny nové databáze a textové soubory musí být standardně ukládány v kódování ****UTF-8 bez BOM****.
2. ****HTTP hlavičky:**** Naše webové servery musí vždy odesílat hlavičku

```
`Content-Type: text/html; charset=utf-8`.
```

3. ****API komunikace:**** Veškerá výměna dat mezi mikroslužbami v našem ****[[VPC]]**** probíhá výhradně v UTF-8.

Časté problémy a řešení

Pokud narazíte na špatné zobrazení znaků (např. „“ nebo „Ě;“):

- **Kontrola editoru:** Ověřte, zda váš textový editor (např. VS Code nebo Notepad++) neotevřel soubor v kódování Windows-1250 (CP-1250).
- **Konverze:** Použijte nástroje jako `iconv` na našich Linuxových serverech pro hromadný převod souborů do UTF-8.
- **BOM (Byte Order Mark):** Vyhýbejte se používání BOM na začátku souborů, protože může způsobovat chyby v automatizovaných skriptech našeho **vývojového týmu**.

Poznámka pro vývojáře: Při práci s databázemi vždy nastavte `character_set_server` na `utf8mb4`, což je vylepšená verze UTF-8 podporující i nejnovější sady emoji a historických symbolů.

— **Související stránky:** [ZIF](#), [WWW](#), [VPS](#), [VPC](#), [VCS](#), [Vývojový tým](#), [IT Podpora](#), [Marketingové oddělení](#)

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

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

Last update: **2026/01/01 16:14**

