

Publikování dat RDF

prof. Ing. Vojtěch Svátek, Dr.

Zimní semestr 2019

<http://nb.vse.cz/~svatek/rzzw.html>

Role publikování ve světě RDF

- Na rozdíl od jiných grafových databázových modelů (např. LPG – labeled property graph) je model RDF primárně určen pro sdílení dat mezi relativně **nezávislými** subjekty
 - Otevřené sdílení na webu: linked open data
 - Sdílení mezi součástmi velké organizace nebo obchodními partnery: linked closed data
- Proto v životním cyklu RDF hraje velkou roli předání dat od jednoho subjektu k druhému, který typicky není předem explicitně specifikován - **publikování**

2 pohledy na publikování RDF

- Producentický: poslední fáze procesu ETL
 - Namísto „natažení“ do uzavřeného datového skladu musí být data zpřístupněna (zcela veřejně nebo v rámci uzavřené, ale široké skupiny „konzumentů“)
 - Viz [Kap. 5 v LD book](#) (nutno rozšířit o nové způsoby)
- Konzumentický: výchozí vstup pro tvorbu LD aplikací nebo vlastní datové analýzy

Současné způsoby publikování

- Samostatný soubor („dump“) ke stažení
- Integrace do jiných dokumentů („snippets“)
 - JSON LD, RDFa, ev. konverze do mikrodat npod.
- „RDF documents“ – navigace po datovém webu, s negociací obsahu
- SPARQL endpoint
- Triple Pattern Fragments
 - V různých variacích – vyplňuje prostor možných modelů mezi dumpem a SPARQL endpointem; vyvážení nároků na inteligenci/zátěž na straně serveru vs. klienta
 - Viz <http://linkeddatafragments.org/>
- Všechny způsoby mají výhody a nevýhody...
 - Nově navržena možnost federovaného dotazování do zdrojů publikovaných v odlišných modelech: <https://comunica.linkeddatafragments.org/>