

Termíny z TDKIV ve Wikidatech: praktické zkušenosti

Linda Jansová

Národní knihovna ČR, Knihovnický institut

linda.jansova@nkp.cz

Knihovny současnosti 2019, 10.–12. září 2019, Olomouc

Pár otázek na úvod

- Znáte Českou terminologickou databázi knihovnictví a informační vědy (TDKIV)?
- Používáte ji?
- Kdo z Vás ji použil od začátku letošního září?
- Odhadnete, kolik hesel obsahuje?

Počet hesel a vazeb mezi nimi

- **přes 3 200 hesel**
- mezi nimi je **necelých 9 000 příbuzenských vazeb**
 - vzájemné vazby počítány jako dvě vazby
- některá hesla bez vazeb, jiná s vysokým počtem vazeb
- průměrně necelé tři příbuzenské vazby na heslo
- příbuzenské vazby jsou velmi obecné, není z nich patrný např. vztah nadřazenosti a podřazenosti mezi hesly

Co s tím?

- byl připraven **Pilotní projekt zpřístupnění TDKIV v podobě propojených dat**
- projekt realizovala Národní knihovna ČR v roce 2018
- byly využity finanční prostředky Ministerstva kultury (dotační program VISK 1)
- **cíle:**
 - zvýšit vzdělávací a informační potenciál České terminologické databáze knihovnictví a informační vědy (TDKIV) prostřednictvím propojených dat
 - zpřesnit vzájemné vztahy mezi hesly

Proč právě Wikidata?

- svobodná databáze všeho s téměř 60 miliony položek
- snadná možnost začlenit svá data do sémantického webu (existující systém, není třeba vlastní softwarové řešení)
- lepší viditelnost dat na webu
- trend v oblasti paměťových institucí
- využitelnost dat nejen v oblasti knihovnictví

Co je cílem Wikidat?

„Wikidata (...) aims to (eventually) **represent everything in the universe in a way computers can understand.**“

zdroj:

<https://www.wired.com/story/inside-the-alex-friently-world-of-wikidata/>

(článek z 18. února 2019)

Jak jsou Wikidata důležitá z pohledu knihoven?

„Did you know that **ARL** (Association of Research Libraries), **IFLA** (International Federation of Library Associations and Institutions), **the PCC** (Program for Cooperative Cataloging), and others **have all identified Wikidata and Wikibase as being of strategic importance to librarians in the coming years?** And that Wikidata provides a platform for museum and library professionals to make works and individuals **more discoverable through authorities and identifiers?**“

zdroj:

<https://listserv.loc.gov/cgi-bin/wa?A2=ind1908&L=MODS&P=2859>

(zpráva z 13. srpna 2019)

Hlavní výstup z projektu

- Metodika pro zpřesňování vztahů mezi hesly v databázi TDKIV a zpřístupňování údajů v systému Wikidata
- dostupná z webu TDKIV:
https://tdkiv.nkp.cz/docs/metodika_propojena_data_verze_1_0

Metodika pro zpřesňování vztahů mezi hesly v databázi TDKIV a zpřístupňování údajů v systému Wikidata

Výstup z Pilotního projektu zpřístupnění TDKIV v podobě propojených dat

Verze 1.0

Zpracovala PhDr. Linda Jansová, Ph.D.
Praha, prosinec 2018

Jak spolupráce s Wikidaty probíhá?

- přidávání/mapování položek, uvádění identifikátorů z TDKIV
- využívání vlastnosti identifikátor termínu v TDKIV:
<https://www.wikidata.org/wiki/Property:P5398>
- přidávání krátkých popisů (*descriptions*) s využitím výkladu v TDKIV jako základu
 - smyslem popisů je rozlišit od sebe jednotlivé položky se stejnými štítky (*labels*)
- přidávání českých a anglických ekvivalentů
- přidávání zpřesněných vztahů

Propojení mezi Wikidaty a TDKIV

item	itemLabel	tdkiv
Q68	počítač	000000049
Q75	internet	000000536
Q978	mem	000002781
Q5290	monitor	000000043
Q4006	mapa	000000997
Q65768732	ilegální knihovna	000001986
Q5300	centrální procesorová jednotka	000013944
Q20139	světelné pero	000000064
Q189900	informační bezpečnost	000000074
Q11661	informační technologie	000000075
Q8811	HyperText Markup Language	000000613
Q12132	kalendář	000002217
Q8209	cyrilice	000002284
Q19652	volné dílo	000002754

zdroj:

<http://tinyurl.com/ybzmyukd>

Propojení mezi TDKIV a Wikidaty

- nástroj Hub: <https://tools.wmflabs.org/hub/>
- umožňuje využití vlastních identifikátorů
- příklad URL: <https://tools.wmflabs.org/hub/P5398:000000543?s=wd>

Termín	 metadata
Termín anglicky	metadata
Výklad termínu	Strukturovaná data, která nesou informace o primárních datech. Pojem metadat je používán především v souvislosti s elektronickými zdroji a vztahuje se k datům v nejšířím smyslu slova (datové soubory, textové informace, obrazové informace, hudba aj.). Funkce metadat je popisná, selekční a archivační. V souvislosti s těmito funkcemi se rozlišují metadata pro účely popisu, správy, právních nároků, technické funkčnosti, užití a archivace. Údaje se obvykle vkládají přímo do zdroje (umísťují se např. v záhlaví dokumentu HTML). [SKLENÁK-2001A:331-335] [GLOSS-DCMI] [GLOSS] [VOJTÁŠEK-2002]
Normativní výklad	Data, která jsou začleněna do ucelené jednotky nebo se vztahují k ucelené jednotce jejím umožňují její výběr. [ČSN ISO 8459-5]
Zdroj výkladu	 SKLENÁK-2001A:331-335 GLOSS-DCMI GLOSS VOJTÁŠEK-2002
Zdroj norm.výkladu	 ČSN ISO 8459-5
Příbuzný termín	 Dublin Core TEI administrativní metadata archivační metadata bibliografická informace datový slovník federativní vyhledávání hlavičkové údaje literál metadata pro právní nároky metadata pro užití metainformace metainformační systém metajazyk popisná metadata registr metadat sekundární informace skrytý text technická metadata
Další informace	 Wikipedie / Wikidata Wikidata
Autor hesla	Celbová, Ludmila
Lektor hesla	Tkačiková, Daniela
Redaktor hesla	Schwarz, Josef Burgetová, Jarmila
Systém. číslo	000000543



[Main page](#)
[Community portal](#)
[Project chat](#)
[Create a new Item](#)
[Create a new Lexeme](#)
[Recent changes](#)
[Random Item](#)
[Query Service](#)
[Nearby](#)
[Help](#)
[Donate](#)

[Print/export](#)

[Create a book](#)
[Download as PDF](#)
[Printable version](#)

Heslo metadata jako příklad

English Not logged in Talk Contributions Create account Log in

Item [Discussion](#) Read [View history](#) Search Wikidata

metadata (Q180160)

data about data [edit](#)
metacontent

[In more languages](#) [Configure](#)

Language	Label	Description	Also known as
English	metadata	data about data	metacontent
Czech	metadata	data o datech	
Slovak	Metadáta	No description defined	
German	Metadaten	Daten über Daten	Metainformation Metadata Metainformationen

[All entered languages](#)

TDKIV term ID

000000543

[edit](#)

[- 0 references](#)

[+ add reference](#)

[+ add value](#)

12

Příklad zpřesněných vazeb

- ▶ |||| extrakce informací 000000453
 - |||| získávání informací
 - |||| informační analýza
- ▶ |||| fakt 000003165
 - |||| data
- ▶ |||| faktografická databáze 000000394
 - |||| databáze
- ▶ |||| faktografický selekční jazyk 000001527
 - |||| selekční jazyk
- ▶ |||| fasetová klasifikace 000001529
 - |||| klasifikační schéma
- ▶ |||| figurální iniciála 000002306
 - |||| iniciála
- ▶ |||| fiktonym 000001251
 - |||| pseudonym
- ▶ |||| filmografie 000000317
 - |||| bibliografie
- ▶ |||| firemní informace 000000396
 - |||| informace

zdroj: <https://w.wiki/86V>

Co spolupráce s Wikidaty přináší

- online **nástroj pro vyjádření přesných vazeb** (subclass of, instance of, part of, opposite of apod.)
- možnost **čerpat inspiraci** z propojení položek ve Wikidatech na další pořadací systémy (např. tezaury) a zdroje jiného typu (např. hesla z různých encyklopedií)
- obecně **širší kontext** (začlenění položek do celého univerza lidského poznání, jak je v daný okamžik zachyceno ve Wikidatech)
- **zdroj** dalších anglických ekvivalentů termínů a dalších **dílčích námětů** k zapracování do TDKIV
- **propojení** hesel přímo v databázi **na položky z Wikidat i hesla (články) ve Wikipedii** (přes 700 hesel)
 - prioritně podle jazyka uživatele (Accept-Language v hlavičce HTTP posílané webovým prohlížečem)

Výhledy do budoucna

- využití zpřesněných příbuzenských vazeb přímo v TDKIV
- vizualizace vztahů pomocí webové aplikace (díky Wikidatům bude možné pracovat s aktuálními daty), která naváže na pracovní aplikaci využívající data v podobě tabulek a XML (na Google Drive)
- zjednodušení procesu zařazování nových hesel do databáze (členové redakční rady TDKIV získají nástroj umožňující pohled na bázi jako celek)
- ...

Děkuji za pozornost!