

# NOVÝ PORTÁL NÁRODNÍ KNIHOVNY PRO NÁRODNÍ AUTORITY

*Marie Balíková – Zdeněk Bartl – Jaroslava Svobodová, Národní knihovna ČR*

Od září letošního roku přibude do rodiny portálů provozovaných Národní knihovnou<sup>1</sup> další přírůstek – portál pro národní autority. Tento nový portál soustředí na jednom místě zatím mnohdy roztržštěné informace o problematice (v nejširším slova smyslu) tvorby a využívání souborů národních autorit – jmenných i věcných.

Problematika tvorby a využívání souborů autorit není jistě v prostředí českých knihoven něčím neznámým (projekt „Kooperativní tvorba a využívání souborů národních autorit“ má za sebou již několik let úspěšné realizace). Potřebné informace k problematice souborů autorit našel dosud každý zájemce/uživatel na těchto URL adresách:

- Informace k VISK9 a projektu „Kooperativní tvorba a využívání souborů národních autorit“ jsou zde (<http://visk.nkp.cz/VISK9.htm> )
- Nejobsažnější informace o souborech autorit jsou zde ([http://www.nkp.cz/pages/page.php3?page=fond\\_authority.htm](http://www.nkp.cz/pages/page.php3?page=fond_authority.htm) – jmenné autority; [http://www.nkp.cz/pages/page.php3?page=fond\\_veczprac1.htm](http://www.nkp.cz/pages/page.php3?page=fond_veczprac1.htm) – věcné autority)
- Dílčí informace lze nalézt též na stránkách k problematice katalogizační politiky Národní knihovny ([http://www.nkp.cz/pages/page.php3?page=fond\\_standard1.htm](http://www.nkp.cz/pages/page.php3?page=fond_standard1.htm) )

## **Portál autorit**

Česká terminologická databáze knihovnictví a informační vědy rozlišuje termíny „portál“ a „vertikální portál“. Termín „portál“ definuje jako „webové sídlo, které poskytuje širokou škálu služeb a informací, často s možností jejich přizpůsobení uživateli podle osobních potřeb a zájmů. Nabídka zahrnuje především vyhledávání webových informačních zdrojů a dokumentů, freemailovou službu, denní zpravodajství, mapy, vyhledávání kontaktních informací na osoby i firmy, zábavu (hry, horoskopy, vtipy

---

<sup>1</sup> JIB, SK CASLIN, Oborová brána „Knihovnictví a informační věda“, Digitální knihovna a další.

apod.), elektronické obchody, bezplatný prostor pro publikování webových stránek apod. Naprostá většina nabízených služeb je pro koncového uživatele bezplatná díky ziskům z reklamy. Široce (obecně) zaměřené portály se označují jako horizontální portály, na rozdíl od vertikálních portálů. Termín „vertikální portál (vortál)“<sup>2</sup>, definuje jako „Specializovaný portál zpřístupňující informační zdroje zaměřené na určitou cílovou skupinu uživatelů, která může být vymezena např. geograficky nebo tematicky“<sup>2</sup>

Portál souboru autorit řadíme do druhé skupiny, neboť je zaměřen na cílovou skupinu uživatelů a jsou zde zpřístupněny veškeré materiály týkající se souboru národních autorit, což, jak předpokládáme, bude pro všechny tvůrce i uživatele velmi užitečné.

### Soubory autorit

V souvislosti s tvorbou portálu pro národní autority vyvstala před námi série otázek: Jaké je místo a úloha souboru autorit v současném informačním procesu? Je přínosný a užitečný? Pro koho: pro indexátory, katalogizátory, uživatele?

Soubor autorit je možno definovat jako soubor ověřených a unifikovaných jmenových a/nebo věcných selekčních údajů, určených pro zpracování a vyhledávání dokumentů s nezbytným odkazovým a poznámkovým aparátem.<sup>3</sup> Soubor autorit představuje národní standard pro jmenové a věcné pořádání a vyhledávání informací, který respektuje mezinárodní standardy a je určen zejména pro databáze s rozsáhlým univerzálním fondem a slouží především jako

- nástroj pro autoritní kontrolu, tedy jako
- nástroj pro standardizaci a unifikaci věcných selekčních údajů na národní úrovni,
- podpůrný nástroj pro sdílenou katalogizaci,
- nástroj pro integraci a unifikaci věcných selekčních jazyků různých typů,
- podpůrný nástroj pro tvorbu vícejazyčného pořadacího systému
- uživatelsky vstřícný rešeršní nástroj pokrývající heterogenní prostředí.

---

<sup>2</sup> Česká terminologická databáze knihovnictví a informační vědy – TDKIV. Dostupné z [www: http://sigma.nkp.cz:4505/F/?func=file&file\\_name=find-a&local\\_base=ktl](http://sigma.nkp.cz:4505/F/?func=file&file_name=find-a&local_base=ktl) , hesla *portál*, *vertikální portál*

<sup>3</sup> Česká terminologická databáze knihovnictví a informační vědy – TDKIV. Dostupné z [www: http://sigma.nkp.cz:4505/F/?func=file&file\\_name=find-a&local\\_base=ktl](http://sigma.nkp.cz:4505/F/?func=file&file_name=find-a&local_base=ktl) , heslo *soubor autorit*

- Z hlediska uživatele je soubor jmenných a věcných autorit koncipován jako nástroj usnadňující proces vyhledávání obsahových informací.

Z hlediska autorů a tvůrců obsahu je soubor jmenných a věcných autorit koncipován jako zdroj ověřování používaných selekčních prvků. Jde o jakousi obdobu bibliografické a autoritní kontroly, nástroje, který bývá již tradičně uplatňován v procesu zpracování a zpřístupnění dokumentů v knihovnách všech typů – tradičních, hybridních i digitálních.

### **Bibliografická a autoritní kontrola**

Cílem bibliografické kontroly je zajistit přijatelným způsobem všeobecnou informaci o publikaci vydané v kterékoli zemi, případně o obsahu dokumentu. Zajištění dostupnosti informací o publikaci, případně o obsahu dokumentu je mnohdy komplikováno existencí různých variantních forem termínů používaných v bibliografickém a metadatovém popisu dokumentů i v samotném plném textu dokumentů. Pro tutéž entitu, např. autor, název díla, nebo entitu označující předmět dokumentu, kterým může být obecné téma (abstraktní termín nebo termín označující objekt), geografický termín, chronologický údaj, a dále pro údaj týkající se žánru, případně formy dokumentu, existuje několik variantních forem (možností označení), což působí potíže při vyhledávání. Odstranění těchto komplikací při vyhledávání informací o dokumentech, případně informací samotných, je cílem autoritní kontroly, která zajistí, že např. všechny varianty téhož díla jsou uloženy pod standardizovanou formou názvu, že tentýž autor bude identifikován jednou formou jména (ostatní varianty budou na tuto preferovanou formu odkázány), že pro tentýž pojem, objekt, místo, událost bude důsledně používán jeden termín a ostatní budou opět na tuto preferovanou formu odkázány.

Běžný uživatel se s výsledky bibliografické a autoritní kontroly, tedy standardizovanými údaji/selekčními prvky, setkává nejenom při vyhledávání v katalozích a databázích sekundárních dokumentů, ale i při vyhledávání v plnotextových a faktografických databázích.

Na příkladu Wikipedie si můžeme demonstrovat potřebnost ověřitelnosti údajů v oblasti terminologie, tedy „autoritní kontroly“ i v současném elektronickém prostředí. Wikipedie je „projekt, jehož jediným podstatným cílem je tvorba encyklopedie; tomu jsou podřízena všechna pravidla. Základními pravidly jsou: nezaujatý úhel pohledu (velmi často nazýván anglickou zkratkou NPOV podle neutral point of view), žádný původní výzkum, ověřitelnost a svobodný obsah... Pravidlo ověřitelnosti požaduje u všech informací odkaz na důvěryhodný zdroj této informace. Informace, u kterých nelze uvést ověřitelný zdroj, ve Wikipedii nemohou být uvedeny...“ (Ka-

dlec 2006)<sup>4</sup>. A právě jedním ze zdrojů ověřitelnosti termínů mohou být soubory jmenných a věcných autorit.

### **Soubory autorit a sémantický web**

Soubory autorit (řízené slovníky) mohou hrát významnou roli i pro konstituování sémantického webu, definovaného terminologickou databází jako „koncept webu založený na obsahu, který je vytvořen a strukturován podle určitých pravidel a standardů a umožňuje tak efektivnější a snadnější vyhledávání informací. Realizace sémantického webu předpokládá implementaci standardů pro sémantickou (RDF), strukturální (XML) a syntaktickou (URI) složku architektury webových dokumentů; výsledkem aplikace uvedených standardů bude konzistentní logická struktura dat, která bude implicitně vyjadřovat význam zaznamenaných informací“<sup>5</sup>.

Jedním ze základních pilířů sémantického webu jsou ontologie, které bývají mnohdy mylně ztotožňovány s tezaury či jinými řízenými slovníky, tedy i se soubory autorit. Zásadní rozdíl mezi ontologiemi a řízenými slovníky popsal výstižně Sklenák:

„Ztotožnění tezauru a ontologie je ... mylné. Lze najít rysy společné, ale i odlišující. Hlavní podobnost spočívá v uspořádání terminologie a v použití stromové struktury. Jak ontologie, tak tezaurus se soustřeďují na terminologické pokrytí určité předmětné oblasti a na vymezení vztahu mezi jednotlivými termíny. Hierarchická struktura slouží k seskupení termínu do kategorií a podkategorií. Ontologie i tezaurus se používají pro popis a organizaci informací. Podstatný rozdíl spočívá v nejednoznačnostech vztahu v tezauru. Tezaurových vztahů (BT, NT, RT, UF apod.) není příliš mnoho a nejsou příliš formálně vymezeny, a proto jsou používány různě... Dobrá ontologie zavádí množství strukturálních i konceptuálních vztahů včetně vztahu nadtřída/podtřída/instance, vztahu k času a dalších podle typu jazyka použitého pro reprezentaci. Obecně obsahuje ontologie mnohem více vztahů, které jsou formálně definovány a jsou jednoznačné. Rozdílné jsou rovněž cíle, s nimiž jsou tezaury a ontologie vytvářeny. Tezaurus se snaží zachytit vztahy mezi termíny, zatímco u ontologie jde o definování konceptu a vztahu mezi nimi. V nejčistší podobě nejde u ontologie o termíny, ale pouze o koncepty, které jsou reprezentovány ve formě nezávislé na termínech libovol-

---

<sup>4</sup> Kadlec, P., *Wikipedie – otevřená encyklopedie*. INFORUM 2006. Dostupné z [www: http://www.inforum.cz/inforum2006/pdf/Kadlec\\_Petr.pdf](http://www.inforum.cz/inforum2006/pdf/Kadlec_Petr.pdf)

<sup>5</sup> Česká terminologická databáze knihovnictví a informační vědy – TDKIV. Dostupné z [www: http://sigma.nkp.cz:4505/F/?func=file&file\\_name=find-a&local\\_base=ktl](http://sigma.nkp.cz:4505/F/?func=file&file_name=find-a&local_base=ktl), heslo *semantický web*

ného přirozeného jazyka. Tezaurus se nesnaží definovat, ale pouze formálně reprezentovat význam konceptu, soustřeďuje se na vztahy mezi termíny v určitém přirozeném jazyce (nebo více jazycích, jde-li o vícejazyčný tezaurs)...Nicméně i u ontologií je nutno použít termíny některého přirozeného jazyka, a proto musí ontologie zahrnovat mapování mezi termíny a koncepty...Rozdíl mezi ontologií a tezauzem vynikne, budeme-li uvažovat využití při vyhledávání. Protože ontologie obsahuje strojově interpretovatelnou definici konceptu, může podporovat terminologické odvozování. Tím se myslí, že dotaz uživatele může být analyzován z hlediska významu použitých termínů, a proto může být přesněji mapován vzhledem k informačním zdrojům. Pomocí ontologie lze např. odvodit, že jeden koncept je speciálním případem jiného, a to díky porovnání jejich logických definic. Je-li koncept B speciálním případem konceptu A, pak může být automaticky klasifikován jako podřízený konceptu A. To dává širší možnosti pro zpracování dotazu a vyhledávání.<sup>66</sup>

Ztotožňovat soubory autorit, řízené slovníky a tezaury s ontologiemi tedy nelze. Přesto jsou soubory autorit pro tvorbu ontologií potřebné. Tvorba ontologií pro sémantický web (zvláště univerzálních) je totiž proces velmi náročný a praxe ukázala, že je problematické, nebo přinejmenším velmi obtížné je vytvářet na „zelené louce“. Při tvorbě ontologií, konkrétně např. pro ověřování termínů, stanovení hierarchické struktury, je žádoucí využít existující zdroje, jako jsou řízené slovníky, tezaury, soubory autorit a klasifikační systémy.<sup>7</sup>

### **Význam souboru autorit**

Odpověď na sérii otázek po smyslu existence souboru autorit v současném informačním procesu je tedy pozitivní. I za současného rozvoje síťových technologií, komunikačních protokolů, kognitivních sémantických technologií (pojmového vyhledávání, které se stává dominantním prvkem sémantického webu) je žádoucí vytvářet, rozšiřovat, spravovat a udržovat autoritní soubory. Tyto zdroje je potřeba aktivně rozvíjet i v době tzv. "googlizace" informačního procesu, kdy zaznamenáváme „obrovský úspěch vyhledávače Google podpořený následně jeho dravou expanzí do mnoha dalších informačních oblastí (digitalizace, obrázky, mapy, ...)“, protože jsou

---

<sup>66</sup> Sklenák, V., *Sémantický web* [online]. INFORUM 2003. Dostupné z WWW: [http://www.inforum.cz/inforum2003/prispevky/Sklenak\\_Vilem.pdf](http://www.inforum.cz/inforum2003/prispevky/Sklenak_Vilem.pdf).

<sup>7</sup> Svátek, V., *Ontologie a WWW*. Dostupné z www: <http://www.datakon.cz/datakon02/>

„krátkozraké představy o tom, že Google vyřeší všechny problémy za nás – že ... stačí si prostě jen počkat, až to Google udělá.“<sup>8</sup>

Soubory autorit – slovníky ověřených a unifikovaných záhlaví selekčních prvků všech typů – i v dnešní době

- jsou nezbytným nástrojem bibliografické a autoritní kontroly (ověřování a sjednocování údajů),
- vytvářejí tak prostředí pro sdílení standardizovaných dat v heterogenním prostředí, a to v národním i celosvětovém měřítku,
- usnadňují proces vyhledávání a umožňují navigaci v rozsáhlých databázích,
- mohou sloužit jako platforma pro tvorbu ontologií a přispět tak k rozvoji sémantického webu.

### Názvové autority

„Novinkou“ v našich národních autoritách jsou názvové autority, tj. autority pro unifikované názvy a autority typu autor/název. Unifikované názvy se vytváří důsledně podle pravidel AACR2R. Slouží především k propojení záznamů pro dílo, které se vyskytuje v různých podobách a pod různými názvy či k rozlišení děl s totožnými hlavními názvy.

Názvové autority se používají – stejně jako jiné autority – jak ve jmenném, tak ve věcném popisu. U záhlaví typu autor/název v bibliografickém záznamu je potřeba vždy dobře zvážit, zda má být zároveň zapsána i autorita – autorita se vytváří pouze v případech, kdy je jako název zapsán unifikovaný název (starověká a středověká díla, díla klasiků, ...). Jmennou část v této autoritě vždy tvoří celá autoritní podoba jména z báze jmenných autorit, tj. včetně všech doplňků a dat.

Na konci r. 2006 byl spuštěn v Národní knihovně zkušební provoz tvorby a využívání názvových autorit pro českou knižní produkci a speciální typy dokumentů (převážně pro hudební díla). Tento typ autorit nespravuje oddělení národních jmenných autorit, ale přímo jednotlivá katalogizační oddělení Národní knihovny (jmenný popis, věcný popis, speciální typy dokumentů, hudební oddělení). Názvovým autoritám nejsou přidělována identifikační čísla.

Příklady záznamů unifikovaných názvů<sup>9</sup>

---

<sup>8</sup> Bartošek, M., *Technologie digitálních knihoven*. INFORUM 2006. Dostupné z WWW: [http://www.inforum.cz/inforum2006/pdf/Bartošek\\_Miroslav.pdf](http://www.inforum.cz/inforum2006/pdf/Bartošek_Miroslav.pdf)

<sup>9</sup> 1) UN-základní včetně odkazů

<b>LDR</b>	-----nz--a22-----n--4500
<b>FMT</b>	UN
<b>001</b>	unn2007380996
<b>003</b>	CZ-PrNK
<b>005</b>	20070328180855.0
<b>008</b>	070111 n acnnaabn-----n-a a-----
<b>040</b>	a ABA001  b cze  d ABA001
<b>130 0</b>	a Tisíc a jedna noc
<b>430 0</b>	a Pohádky tisíce a jedné noci  0 o
<b>430 0</b>	a Alf laylah wa-laylah  0 o
<b>430 0</b>	a Alf laila wa-laila  0 o
<b>430 0</b>	a Thousand and one nights  0 o
<b>430 0</b>	a Book of the thousand nights and one night  0 o
<b>430 0</b>	a Arabian nights  0 o
<b>430 0</b>	a Tausendundeine Nacht  0 o
<b>430 0</b>	a Mille et une nuits  0 o
<b>430 0</b>	a Mille e una notte  0 o
<b>670</b>	a The Arabian nights and orientalism: perspectives from East & West /Yuriko Yamanaka and Tetsuo Nishio
<b>670</b>	a LC Titles
<b>906</b>	a za20070111  b java
<b>906</b>	a fr20070328  b maba

## 2) osobní jméno/název

<b>LDR</b>	-----nz--a22-----n--4500
<b>FMT</b>	JN
<b>001</b>	aun2007400029
<b>003</b>	CZ-PrNK
<b>005</b>	20070607162606.0
<b>008</b>	070607 n acnnaabn-----n-a a-----
<b>040</b>	a ABA001  b cze
<b>1000</b>	a Augustin,  c svatý,  d 354-430.  t De civitate Dei
<b>4000</b>	a Augustin,  c svatý,  d 354-430.  t O obci boží  0 o
<b>906</b>	a br20070607  b jvasv

a pro hudební dílo

<b>LDR</b>	-----nz--a22-----n--4500
<b>FMT</b>	JN
<b>001</b>	aun2007401616
<b>003</b>	CZ-PrNK
<b>005</b>	20070614150611.0
<b>008</b>	070614 n acnnaabn-----n-a a-----
<b>040</b>	a ABA001  b cze
<b>10012</b>	a Piccinini, Alessandro,  d 1566-ca 1638.  t Intavolatura di liuto  n (1623).  p Partite variate sopra la folia aria romanesca
<b>4001</b>	a Piccinini, Alessandro,  d 1566-ca 1638.  t Partite variate sopra la folia aria romanesca  0 o
<b>4001</b>	a Piccinini, Alessandro,  d 1566-ca 1638.  t Aria romanesca  0 o
<b>4001</b>	a Piccinini, Alessandro,  d 1566-ca 1638.  t Partite sopra la folia aria romanesca  0 o
<b>670</b>	a Intavolatura di Tiorba, p2006  b (Piccinini A. Intavolatura di liuto et di chitarrone, Libro primo, Bologna 1623 (Partite sopra la folia aria Romanesca))
<b>670</b>	a LC (Names)
<b>906</b>	a br20070614  b jast

3) korporace/název (nejčastěji se používá u zákonů a výročních zpráv)

<b>LDR</b>	-----nz--a22-----n--4500
<b>FMT</b>	KN
<b>001</b>	kon2006374614
<b>003</b>	CZ-PrNK
<b>005</b>	20061124161953.0
<b>008</b>	061124 n acnnaabn-----n-a a-----
<b>040</b>	a ABA001  b cze
<b>1102</b>	a Česká národní banka.  t Výroční zpráva (Česká národní banka).  l Anglicky
<b>670</b>	a NKC
<b>906</b>	a im20061124



## **Závěr**

Nový portál pro národní autority nepřinese tedy v zásadě žádné nové informace. Nabídne však zájemcům/uživatelům komfort, který charakterizuje webový portál, tj. poskytne širokou škálu služeb a informací specificky zaměřeným/orientovaným uživatelům. Těmi jsou jak pracovníci knihoven, tak i uživatelé vně těchto informačních institucí. Knihovníci nepochybně uvítají možnosti nalézt vše potřebné z celé široké škály informací k dané problematice – od odkazů na české interpretace (k jednotlivým druhům autorit) katalogizačních pravidel AACR2, přes informace o komunikačních formátech MARC21, resp. UNIMARC, samotnou metodiku vytváření autoritních záznamů všech u nás používaných druhů autorit, až po informace, jak se např. zapojit do kooperativního projektu Tvorba a využívání souborů národních autorit, příp se zapojit do spolupráce na tvorbě věcných autorit, nebo dovědět se něco o průběhu řešení projektu Národní autority v prostředí muzeí a galerií – interoperabilita s NK ČR...

Novým portálem pro národní autority nabídne Národní knihovna další „přidanou hodnotu“ všem knihovnám, které tuto hledají...