

# Znalostní báze pro obor organizace znalostí

Helena Kučerová, Eva Bratková

Ústav informačních studií a knihovnictví, FF UK Praha, U Kříže 8, 158 00 Praha 5  
kucerova@sks.cz, eva.bratkova@ff.cuni.cz

**Abstrakt.** Příspěvek představuje výsledky prvního roku řešení výzkumného projektu Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI) “Znalostní báze pro obor organizace informací a znalostí”, řešeného na ÚISK FF UK v letech 2013–2015. Cílem projektu je vytvořit znalostní bázi, která má zaplnit mezeru v domácí odborné literatuře tím, že shromáždí a systemizuje aktuální poznatky z oboru organizace znalostí a umožní jejich zpřístupnění a odvozování znalostí nových. Znalostní báze poskytne zdrojový materiál pro zpracování původní české monografie a pro aktualizaci české terminologie. Bude využita jako zdroj pro výuku a vzdělávání především v oboru informační vědy a knihovnictví.

**Klíčová slova:** organizace znalostí, systém organizace znalostí, informační věda, informační zdroje, znalostní báze

## 1 Úvod

Obor organizace znalostí prochází v současné době změnou paradigmatu, vyvolanou používáním elektronických dokumentů a technologií Internetu. Oblast zkoumání se rozšiřuje za hranice tradičních paměťových institucí [2], pojmová a terminologická základna se redefinuje a není jednotná [1], [5]. V rámci řešeného projektu používáme dále uvedené vymezení organizace znalostí jako vědního oboru, organizace znalostí jako praktické aktivity, a systému organizace znalostí. Předmětem zkoumání oboru organizace znalostí je proces organizace znalostí a jeho kontext, tj. zdroje, jež jsou v procesu organizace transformovány, a produkty, jež jsou tímto procesem vytvářeny, včetně zúčastněných aktérů – osob, institucí, technologií. Procesem organizace znalostí rozumíme jakoukoli záměrnou činnost, spočívající v zavádění struktury do existujících znalostí s cílem usnadnit k nim přístup. Tyto aktivity zasahují takřka celý životní cyklus znalostí, od jejich reprezentace přes případné sdělování či sdílení, zaznamenání, publikování, zpracování (identifikaci, popis, obsahovou analýzu, indexaci, transformaci) až po uložení. Jednotkou organizace je zaznamenaná znalost. Výstupem procesu organizace znalostí je artefakt označovaný jako struktura, a obvykle i metadata o organizovaných entitách. Dalším artefaktem jsou systémy organizace znalostí [3], např. slovníky, seznamy autorit, předmětové hesláře, klasifikace, tezaury a ontologie, které jsou současně výstupem procesu organizace znalostí i jeho nástrojem, sloužícím pro organizaci znalostí a zároveň jako pomůcka

při vyhledávání z organizované kolekce. Jádrem každého systému organizace znalostí je slovník a schéma, definující obsah a strukturu organizovaného celku.

## 2 Motivace a cíl projektu

Cílem projektu je shromáždit a systemizovat současný stav poznání v oboru organizace znalostí ve formě znalostní báze, která umožní ukládání, prohlížení, vyhledávání stávajících a odvozování nových znalostí. Znalostní báze bude zpřístupněna ve formátu otevřených propojených dat a bude sloužit pro další výzkum a jako nástroj výuky a vzdělávání především v oboru informační vědy a knihovnictví.

Motivací pro realizaci projektu se staly výsledky průzkumu výchozího stavu poznatkové základny oboru organizace znalostí v českém prostředí, realizovaného v roce 2012. Bylo zjištěno, že zatímco používané technologie jsou srovnatelné se světovým vývojem, chybí aktuální odborná monografie, která by reagovala na změnu paradigmatu organizace znalostí. Vzhledem k závislosti organizace znalostí na přirozeném jazyce a na specifickém kulturním kontextu je absence aktuálních zdrojů v mateřském jazyce pocíťována v rámci české odborné veřejnosti jako citelný nedostatek.

## 3 Použité metody

Volbu metod ovlivnila zvolená forma znalostní báze, která je koncipována jako „tradiční“ expertní systém, jehož obsah je založený na manuální akvizici znalostí a na intelektuálním posuzování jejich relevance a hodnoty. V návrhu znalostní báze se snažíme o maximální opětovné využití již ověřených konceptů a standardizovaných slovníků. Konceptuální model využívá FRBR<sup>1</sup> entity „dílo“, „vyjádření“ a „provedení“ [4] v kombinaci s atributy KOS-AP<sup>2</sup> [6]. Každá jednotka znalostní báze je klasifikována s využitím Klasifikačního systému literatury oboru organizace znalostí<sup>3</sup>. Typologie systémů organizace znalostí je mapována do typologie NKOS<sup>4</sup> Vocabularies<sup>5</sup>. Specifikace požadavků a případy užití znalostní báze byly zformulovány s využitím analýzy provedených řízených rozhovorů s doménovými experty. Dokumentové zdroje byly shromážděny standardními postupy bibliografické rešerše. Znalosti ve formě výroků jsou manuálně excerpovány z 50 vybraných odborných

---

<sup>1</sup> Functional Requirements for Bibliographic Records – funkční požadavky na bibliografické záznamy. Konceptuální datový model bibliografického univerza.

<sup>2</sup> Dublin Core Metadata Application Profile for KOS Vocabulary Registries – aplikační profil Dublinského jádra pro popis systémů organizace znalostí.

<sup>3</sup> *Classification System for Knowledge Organization Literature* [online]. ISKO 2011-2012, last update 2012-04-13. Dostupné z: <http://www.isko.org/scheme.php>

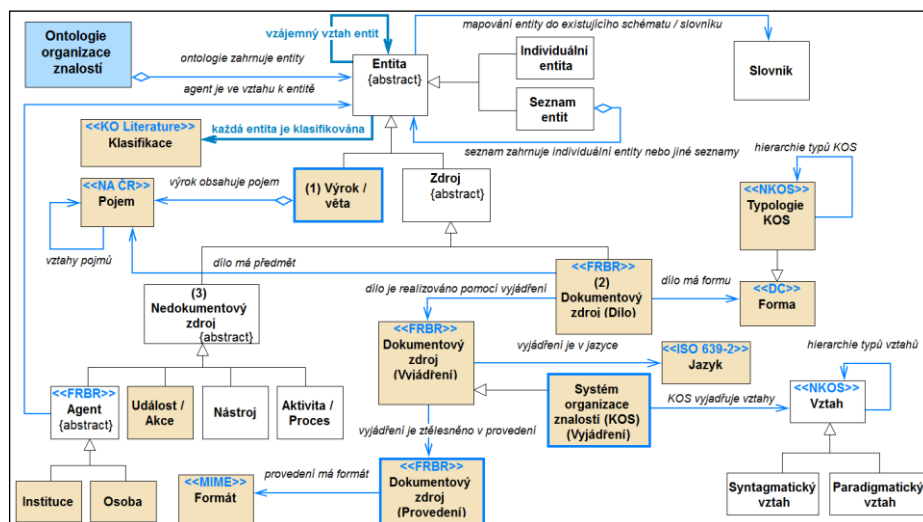
<sup>4</sup> Networked Knowledge Organization Systems/Services – neformální komunita odborníků zajímajících se o využití systémů organizace znalostí v prostředí informačních sítí (<http://nkos.slis.kent.edu>)

<sup>5</sup> NKOS Vocabularies. In: *DCMI NKOS Task Group* [online]. 2013, modified 2013-12-16. Dostupné z: [http://wiki.dublincore.org/index.php/NKOS\\_Vocabularies](http://wiki.dublincore.org/index.php/NKOS_Vocabularies)

textů. K uložení obsahu prototypu znalostní báze je využita databáze MySQL s webovým rozhraním (43 tabulek, 82 atributů), výroky jsou formátovány jako RDF trojice. Pro ověřovací fázi manuální akvizice znalostí plánujeme uplatnit metody automatické extrakce termínů z textu a automatické klasifikace či klastrování.

## 4 Dosavadní výsledky

Byl zpracován konceptuální model struktury znalostní báze ve formátu ontologie (viz obr. 1). Jádrem ontologie jsou tři typy entit: 1. výroky (věty formalizované jako logické predikáty, instanciované formou textových dat), 2. dokumentové zdroje včetně systémů organizace znalostí, reprezentované popisnými metadaty, 3. nedokumentované zdroje (osoby, instituce, události, nástroje, aktivity a procesy), reprezentované rovněž popisnými metadaty.



Obr. 1. Konceptuální model znalostní báze

Proběhla manuální fáze akvizice znalostí. Je připraveno zahájení ověřovací fáze s využitím automatizovaných metod sémantické analýzy shromážděných popisných metadat a vybraných plných textů odborných publikací a jejich rejstříků. Byl vytvořen prototyp znalostní báze s cca 2300 jednotkami deklarativních znalostí: 150 výroků ve formátu RDF, 150 záznamů systémů organizace znalostí, 1000 bibliografických záznamů odborných dokumentů, slovník o rozsahu 1000 termínů.

## 5 Řešené problémy a plány dalšího výzkumu

Problémy, řešené v průběhu dosavadních etap projektu, se týkaly následujících témat: 1. Použitelnost zvoleného klasifikačního systému pro instance zahrnuté ve

znalostní bázi. 2. Typologie systémů organizace znalostí. 3. Aplikace modelu FRBR na dokumentové entity. 4. Použitelnost principu sémantické diverzity, uplatňovaného v sémantickém webu, na doménovou znalostní bázi. 5. Formalizace deklarativních výroků. V průběhu dalšího řešení projektu plánujeme rozšíření typů zahrnutých entit o procedurální znalosti. Další úlohou bude implementace znalostní báze do znalostního systému s inferenčním softwarem a návrh uživatelského rozhraní.

**Poděkování.** Projekt č. DF13P01OVV013 je realizován na Filozofické fakultě Univerzity Karlovy a je financován Ministerstvem kultury ČR v rámci Programu aplikovaného výzkumu a vývoje národní a kulturní identity (NAKI).

### English summary

#### *Knowledge base for subject area of knowledge organization*

The paper presents first year results of a research project “Knowledge base for subject area of knowledge organization”, conducted 2013–2015 within the Programme of applied R&D in National and Cultural Identity (NAKI) on Faculty of Arts of Charles University in Prague. The aim of the project is to design a knowledge base, which will fill in the gap in Czech scholarly literature through assembling and systemizing the current body of knowledge for the domain of knowledge organization and will enable for storing, browsing, and searching of existing knowledge as well as for discovering or inferencing of new knowledge. The content of knowledge base will provide background material for the current scientific monograph. It will serve as a resource for updating the Czech terminology, and as an education and training tool particularly in the field of information science and librarianship.

### Literatura

1. DAHLBERG, I. What is knowledge organization? In: *Knowledge organization*. 2014, 41(1), 85-91. ISSN 0943-7444
2. GLUSHKO, R. J., ed. *The discipline of organizing*. Cambridge(MA): MIT Press, 2013. 475 s. ISBN 978-0-262-51850-5
3. HODGE, G. *Systems of knowledge organization for digital libraries: beyond traditional authority files*. Washington: The Digital Library Federation, Council on Library and Information Resources, 2000. 37 s. ISBN 978-1-933645-06-3
4. *Funkční požadavky na bibliografické záznamy*. Praha: Národní knihovna ČR, 2002. 117 s. ISBN 80-7050-400-5
5. SMIRAGLIA, R. P. Theory, knowledge organization, epistemology, culture. In: SMIRAGLIA, R. P., LEE, H., ed. *Cultural frames of knowledge*. Würzburg: Ergon-Verlag, 2012, Chapter 6, s. 1-18. ISBN 978-3-89913-918-1
6. ZENG, M. L., ŽUMER, M. Metadata Application Profile for KOS Vocabulary Registries. In: *Knowledge Organization: pushing the boundaries: proceedings of the ISKO UK 3rd Biennial Conference, 8th - 9th July 2013, London* [online]. J. Vernau, ed. London: ISKO UK, 2013 [cit. 2014-06-18]. Dostupné z: <http://www.iskouk.org/conf2013/papers/ZengPaper.pdf>