

# Zápis z 11. zasedání Rady pro koordinaci Polytematického strukturovaného hesláře (PSH) ze dne 16. 12. 2010

## Přítomni:

Ing. Martin Svoboda (NTK), PhDr. Milada Píšková (NTK), PhDr. Radka Římanová, PhD. (NTK), Mgr. Kateřina Kamrádková (NTK), Bc. Kristýna Kožuchová (NTK), Bc. Ctibor Škuta (NTK), Ing. Miloslav Nič, PhD. (VŠCHT), Mgr. Linda Jansová, PhD. (AV ČR), Mgr. Štěpánka Žižková (NTK), Bc. Jindřich Mynarz (NTK), Ilona Poláčková (ČVUT), Ing. Lukáš Slánský (Univerzita Pardubice), Mgr. Barbora Ševčíková (MKP), Pavlína Omastová, BBus (Hons) (NTK), Mgr. Lenka Maixnerová (NLK), Mgr. Václav Heicl (NTK)

## Omluveni:

Mgr. Marie Balíková (NK ČR)

Zasedání Rady pro koordinaci Polytematického strukturovaného hesláře (PSH) zahájila Kristýna Kožuchová, pracovnice referátu PSH, která přítomné přivítala a následně prezentovala výsledky činnosti referátu PSH od minulého zasedání Rady pro koordinaci PSH, které se konalo 8. 12. 2009. Dále stručně představila novou verzi PSH 3.0 a přednesla plán na rok 2011.

Činnost referátu PSH v roce 2010 byla shrnuta do dvou hlavních bodů:

- **aktualizace verze PSH 2.1**
  - formulář pro návrh nových hesel PSH
  - výběr nových hesel PSH z pole 653
  - analýza logu vyhledávání
- **propagace**
  - konference
  - semináře
  - časopisy
  - Wikipedie

Referát PSH se v roce 2010 zabýval především aktualizací verze PSH 2.1. Proces aktualizace hesláře je optimalizován několika způsoby.

Jedním z nich je analýza a následný výběr nových hesel PSH z pole 653 v bibliografickém záznamu ve formátu MARC 21, které je používáno při věcném popisu knih v NTK. Toto pole je v NTK používáno pouze doplňkově, a to v případech, kdy danou skutečnost nelze dostatečně přesně vyjádřit prostřednictvím hesel PSH (jedná se o podrobnější úroveň zpracování tématu) nebo se v hesláři odpovídající heslo nevyskytuje (jedná se o nový termín).

Dalším způsobem optimalizace je analýza logu vyhledávání. Ta spočívá v rozboru uživatelských dotazů zadávaných při vyhledávání v katalogu NTK. Log vyhledávání představuje soubor, do kterého se zaznamenávají uživatelské požadavky společně s textovým záznamem komunikace mezi zúčastněnými počítačovými vrstvami a systémovými informacemi. V první fázi analýzy je nutné odfiltrovat nepotřebný text. Tento proces využívá schopností regulárních výrazů, které představují univerzální,

uživatelsky nadefinovanou šablonu struktury textu. Při nalezení části textu shodné s regulárním výrazem je možné tato data extrahovat a výsledky použít pro jednoduchou statistickou analýzu. Cílem algoritmu jsou termíny zadávané při vyhledávání ve všech polích, v poli předmětovém a v heslech PSH. Získané termíny se periodicky přidávají do výstupního souboru. Při vyšší frekvenci některého z nich je heslo pracovníky referátu PSH posuzováno jako kandidát na nové heslo PSH.

Třetím, doplňkovým způsobem získávání nových hesel PSH je webový formulář pro zasílání návrhů nového hesla PSH. Využívat jej mohou jak katalogizátoři oddělení věcného zpracování v NTK, tak externí uživatelé, např. zájemci o problematiku selekčních jazyků, ale i zákazníci NTK, kteří se prostřednictvím formuláře stávají tvůrci samotného obsahu PSH.

Dále byli přítomní seznámeni s propagací PSH v uplynulém roce. Do této kategorie spadají příspěvky z konferencí Inforum (*Polytematický strukturovaný heslář a jeho potenciál v oblasti třídění a zpřístupňování webových dokumentů*), WebExpo (*Národní technická knihovna – za stovku služby za milion*), Knihovny současnosti 2010 (*Automatizace knihovnických procesů s využitím technologie SikuLi*) a 12. ročníku Mezinárodní konference o šedé literatuře (*Integration of an Automatic Indexing System within the Document Flow of a Grey Literature Repository*). V neposlední řadě zde patří také článek publikovaný v časopise Knihovna (*Vliv trendů systémů organizace znalostí na vývoj Polytematického strukturovaného hesláře v Národní technické knihovně*) a příspěvek ze semináře Co se skrývá za vyhledáváním aneb Searching Session NTK 2010 (*Polytematický strukturovaný heslář*). Nově má PSH článek v anglické Wikipedii.

Také bylo konstatováno, že v roce 2010 byl PSH implementován do systému knihovny BIVŠ, a.s.

Poté vystoupil Ctibor Škuta, pracovník referátu PSH, který představil nový PSH Manager. Ten byl vytvořen, aby usnadnil indexaci dokumentů katalogizátorům NTK. PSH Manager je lokálním zdrojem dat, tudíž pracuje nezávisle na připojení do sítě. Rovněž informoval o právě probíhajícím testování využití hesel PSH při automatické indexaci v Národním úložišti šedé literatury (NUŠL).

Následně byl přednesen návrh činností referátu PSH na rok 2011:

- pokračující propagace PSH,
- negativní analýza hesel PSH,
- uspořádání semináře o možnostech využití PSH,
- automatická indexace (NUŠL) s hesly PSH,
- průběžná analýza logů,
- začlenění domény „techlib.cz“ do URI hesel PSH,
- spolupráce s českou Wikipedií,
- zvážení možnosti linkování PSH na další zdroje.

Na závěr vystoupil Jindřich Mynarz, pracovník oddělení rozvoje elektronických služeb, který hovořil o možnostech linkování PSH. Vyjádřil názor, že bude-li nějaký zdroj odkazovat na PSH, zvýší se tím důvěryhodnost PSH.

Po prezentaci byla zahájena volná diskuze, v rámci které byli členové Rady pro koordinaci PSH a přítomní hosté vyzváni, aby reagovali na přednesené příspěvky.

Ing. Lukáš Slánský navrhl, aby byl PSH implementován do fulltextových systémů, např. DSpace. Zároveň informoval o realizaci indexace znalostní báze zaměstnanců na Univerzitě v Pardubicích pomocí hesel PSH.

Mgr. Linda Jansová, PhD., podobně jako Ing. Miloslav Nič, PhD., navrhla vytvořit ukázkovou aplikaci indexace vybraného souboru dokumentů (např. učebnice) a výsledky prezentovat na konferenci nebo v časopise.

Členové Rady pro koordinaci PSH dále přednesli návrh, aby se v rámci marketingu PSH postupovalo v oslovování dalších subjektů mimo knihovnictví (např. právnické firmy) s nabídkou PSH jako indexačního nástroje.

V příštím roce doporučili nabídnout PSH Manager (desktopová aplikace) spolupracujícím knihovnám a dalším institucím. Zároveň bude nutné vyřešit aktualizace PSH v tomto programu.

Z diskuze také vyplynula potřeba definovat si strategické cíle, kterých by referát PSH chtěl v příštích letech dosáhnout.

## ZÁVĚR

Seznam úkolů pro rozvoj PSH v roce 2011 - výsledky práce budou prezentovány na zasedání Rady pro koordinaci PSH, která se bude konat v prosinci 2011:

1. Průběžně informovat o rozvoji PSH:
  - 1.1. internetová prezentace <http://www.techlib.cz/cs/katalogy-a-databaze/psh/>,
  - 1.2. heslo PSH ve Wikipedii - kontrola změn a aktualizace, minimálně 4x ročně,
  - 1.3. články v odborných časopisech pro knihovny,
  - 1.4. prezentace na odborných konferencích,
  - 1.5. e-mailová konference knihovna@cesnet.cz - zaslání minimálně jednoho informativního mailu o aktualitách.
2. Průběžná analýza logu vyhledávání v katalogu NTK.
3. Distribuovat novou verzi PSH 3.0.
4. Uspořádat 2. ročník semináře *Searching Session NTK* pro prezentaci PSH a jeho možností.
5. V prosinci 2011 svolat Radu pro koordinaci PSH.

Rada pro koordinaci PSH souhlasí s plánem činnosti referátu PSH na rok 2011 a doporučuje zohlednit své návrhy na rozvoj PSH.

Četl: Ing. Martin Svoboda