



národní
úložiště
šedé
literatury

Polytematický strukturovaný heslář a jeho potenciál v oblasti třídění a zpřístupňování webových dokumentů

Kožuchová, Kristýna; Škuta, Ctibor
2010

Dostupný z <http://www.nusl.cz/ntk/nusl-41956>

Dílo je chráněno podle autorského zákona č. 121/2000 Sb.

Licence Creative Commons Uveďte autora-Neužívejte dílo komerčně-Zachovejte licenci
3.0 Česko

Tento dokument byl stažen z Národního úložiště šedé literatury (NUŠL).

Datum stažení: 22.09.2019

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní nusl.cz .

NTK

50°6'14.083"N, 14°23'26.365"E

Národní technická knihovna
National Technical Library



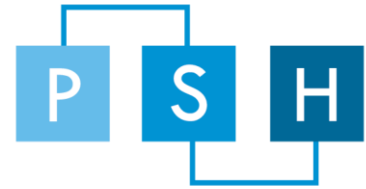
210 mm

Polytematický strukturovaný heslář a jeho potenciál v oblasti třídění a zpřístupňování webových dokumentů

Kristýna Kožuchová, kristyna.kozuchova@techlib.cz

Ctibor Škuta, ctibor.skuta@techlib.cz

INFORUM 2010, Praha, 25.–27. 5. 2010



Obsah

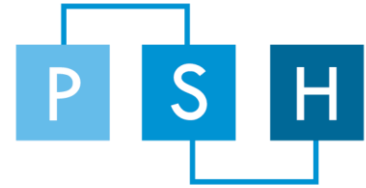
- Systémy organizace znalostí
- Polytematický strukturovaný heslář
- Organizace webového obsahu
- Generování úryvků metadat s hesly PSH
- Folksonomie, tagování
- Formáty PSH
- Další datové zdroje
- Přínosy



Systemy organizace znalostí (Knowledge organisation systems)

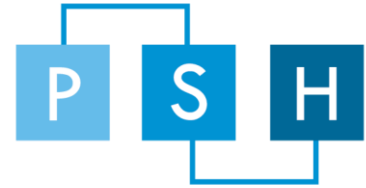
- systémy zajišťující konzistenci předmětové indexace
- taxonomie systémů organizace znalostí

Soubory termínů	Klasifikace	Schémata vztahů
Autoritní záznamy	Předmětová hesla	Tezaury
Slovníky	Klasifikační schémata	Ontologie
	Taxonomie	Sémantické sítě



Polytematický strukturovaný heslář (PSH)

- česko-anglický řízený slovník
- více než 13 600 hesel
- 44 tématických oblastí
- každé heslo je zařazeno v rámci šesti-, výjimečně sedmistupňové hierarchie
- hierarchické, ekvivalenční a asociační vztahy
- PSH/SKOS k dispozici [zdarma ke stažení](#) pod licencí CC



Organizace webového obsahu

- hierarchické třídění
- folksonomie, tagování
- uživatelsky tvořená hesla s definovanou strukturou (SOBOLEO)
- metadatové formáty



Generování úryvků metadat s hesly PSH

- tagy ve formátech: [Dublin Core](#) a [CommonTag](#)

Níže uvedené úryvky metadat lze vložit do HTML dokumentu pro dosažení jeho sémantického popisu:

Dublin Core:

```
<p about="" xmlns:dc="http://purl.org
/dc/elements/1.1/">
  <a href="http://psh.ntkcz.cz/skos/PSH3634"
  rel="dc:subject">jaderná fyzika</a>
</p>
```

CommonTag:

```
<body xmlns:ctag="http://commontag.org/ns#"
rel="ctag:tagged">
  <span typeof="ctag:Tag" rel="ctag:means"
  resource="http://psh.ntkcz.cz/skos/PSH3634"
  property="ctag:label" content="jaderná fyzika"
  />
</body>
```





Folksonomie, tagování

- tagy:
 - jednotlivá klíčová slova
 - vytváří je sami uživatelé
 - [Delicious](#), [Digg](#), [Bookmarky.cz](#), [Amazon](#)
 - možnost využití v knihovnách
 - kombinace PSH a tagování v katalogu NTK



Formáty PSH

- MARC pro autoritní záznamy
- SKOS (Simple Knowledge Organisation System)
 - určený k reprezentaci, sdílení a odkazování znalostních systémů
 - W3C standardy RDF (Resource Description Framework) a RDFS (RDF - Schema)



210 mm

Name Tag	Data
*****	Záznam dokumentu 1 z 1
FMT	\VA
LDR	-----nz--a22-----n--4500
001	PSH3634
003	CZ-PrSTK
005	20070126161437.0
008	070126na ann bab ----- - a a-----
040	a ABA013 b cze
150	a jaderná fyzika x fy
550	a jaderné reaktory x en
550	a jaderná chemie x ch
5501	w h a rozptyl částic x fy
5501	w h a svazky nabitých částic x fy
5501	w h a urychlovače nabitých částic x fy
5501	w h a atomové jádro x fy
5501	w h a fyzika nízkých energií x fy
5501	w h a fyzika vysokých energií x fy
5501	w h a elementární částice x fy
5509	w g a fyzika x fy
75007	a nuclear physics 2 epsh
CAT	c 20070127 STK10 h 0945
SYS	000003635

Záznam hesla PSH ve formátu MARC21 pro autority



```
<rdf:RDF>
- <skos:Concept rdf:about="http://psh.ntkcz.cz/skos/PSH3634">
  <skos:prefLabel xml:lang="cs">jaderná fyzika</skos:prefLabel>
  <skos:prefLabel xml:lang="en">nuclear physics</skos:prefLabel>
  <skos:related rdf:resource="http://psh.ntkcz.cz/skos/PSH2543"/>
  <skos:related rdf:resource="http://psh.ntkcz.cz/skos/PSH5645"/>
  <skos:broader rdf:resource="http://psh.ntkcz.cz/skos/PSH2910"/>
  <skos:narrower rdf:resource="http://psh.ntkcz.cz/skos/PSH3639"/>
  <skos:narrower rdf:resource="http://psh.ntkcz.cz/skos/PSH3694"/>
  <skos:narrower rdf:resource="http://psh.ntkcz.cz/skos/PSH3695"/>
  <skos:narrower rdf:resource="http://psh.ntkcz.cz/skos/PSH3693"/>
  <skos:narrower rdf:resource="http://psh.ntkcz.cz/skos/PSH3649"/>
  <skos:narrower rdf:resource="http://psh.ntkcz.cz/skos/PSH3635"/>
  <skos:narrower rdf:resource="http://psh.ntkcz.cz/skos/PSH3636"/>
  <skos:inScheme rdf:resource="http://psh.ntkcz.cz/skos"/>
</skos:Concept>
</rdf:RDF>
```



MARC x SKOS

MARC	Popis	SKOS
150	Záhlaví - věcné téma	skos:prefLabel
450	Odkaz viz - věcné téma	skos:altLabel
5509	Nadřazený termín - věcné téma	skos:broader
5501	Podřazený termín - věcné téma	skos:narrower
550	Viz též - věcné téma	skos:related



Další datové zdroje

- SKOS je aplikace RDF
- RDF podporuje distribuované zpracování dat
- odkazy na další datové zdroje

- PSH/SKOS odkazuje na:
 - Předmětová hesla Kongresové knihovny (LCSH)
 - DBpedii

NTK

50°6'14.083"N, 14°23'26.365"E

Národní technická knihovna
National Technical Library











210 mm

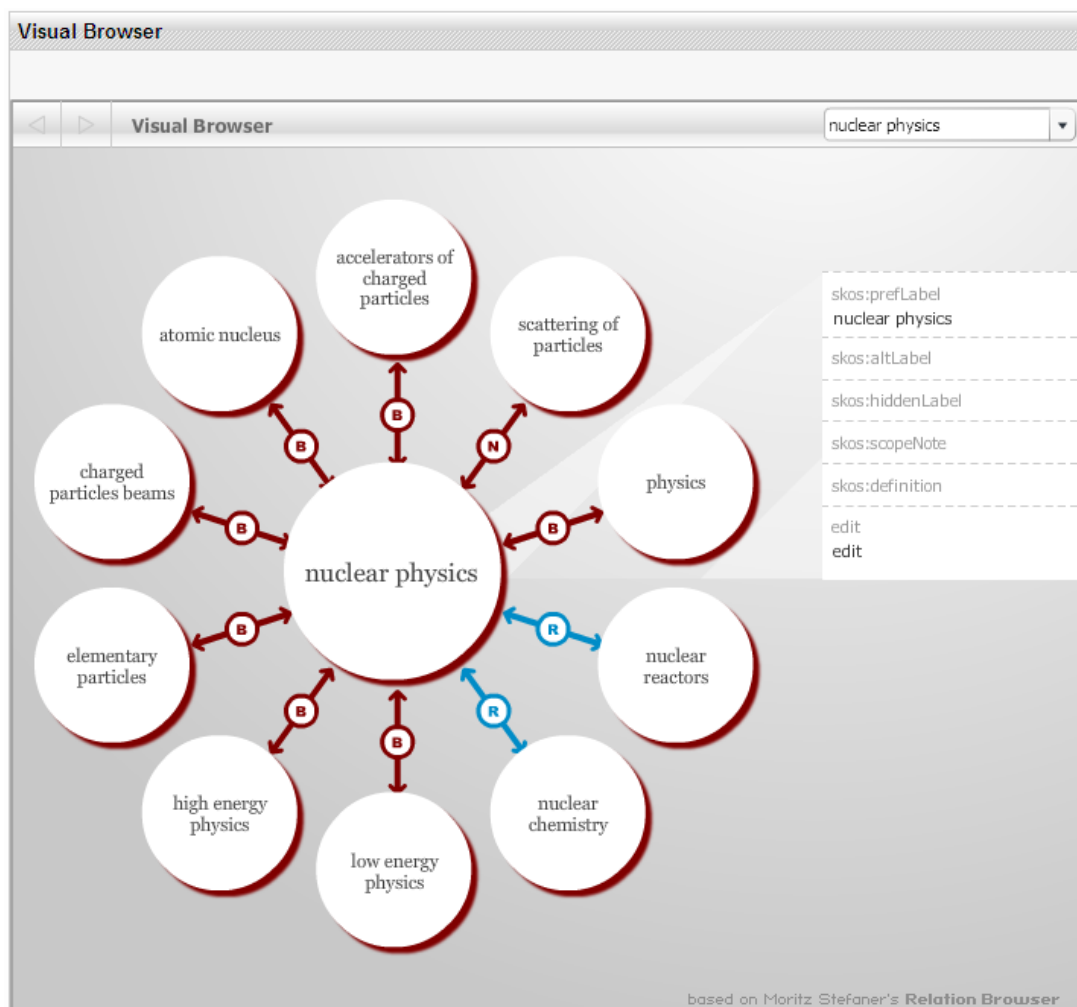
Přínosy

- softwarové nástroje z funkčně odlišných oblastí
 - validátory
 - vizualizační nástroje
 - automatická indexace dokumentů
 - a další...



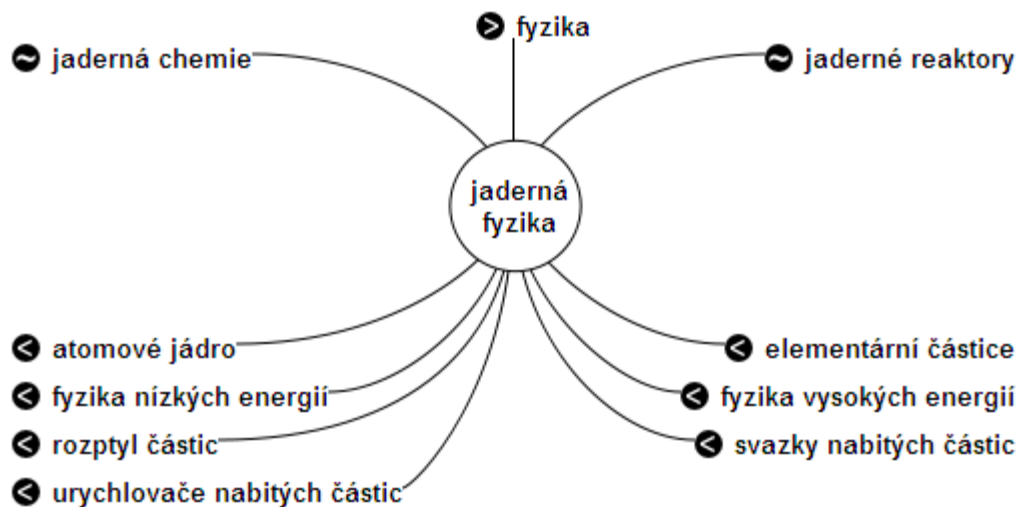
Results

 Valid URIs: Passed! <i>Checks if URIs are valid and do not contain any invalid characters like whitespaces.</i>
 Missing Language Tags: Passed! <i>Handles missing language tags of all sorts of SKOS labels and textual content.</i>
 Missing Labels: ConceptScheme with URI <http://psh.ntkcz.cz/skos/> has no label! <i>Checks for missing labels - prefLabels for skos:Concepts and rdfs:labels for skos:ConceptSchemes.</i>
 Loose Concepts: Found 1 loose concepts. <i>This checks handles loose concepts, i.e. concepts that are no topconcept in any scheme and have no broaders.</i>
 Disjoint OWL Classes: Passed! <i>Checks if there are any instances of owl:Classes that are declared disjoint.</i>
 Consistent Use of Labels: The skos:Concept with the URI <http://psh.ntkcz.cz/skos/PSH7108> has the same skos:prefLabel ("pravděpodobnostní logiky"@cs) and skos:altLabel ("pravděpodobnostní logiky"@cs) <i>Checks if there are concepts with clashing SKOS labels, i.e.:</i> <ul style="list-style-type: none">• More than one prefLabel in the same language• A prefLabel that is also a hiddenLabel• A prefLabel that is also an altLabel• An altLabel that is also a hiddenLabel
 Consistent Usage of Mapping Properties: Passed! <i>Checks if concepts are connected by clashing SKOS mapping relations:</i> <ul style="list-style-type: none">• skos:exactMatch and skos:broadMatch• skos:exactMatch and skos:relatedMatch
 Consistent Usage of Semantic Relations: The skos:Concepts <http://psh.ntkcz.cz/skos/PSH10349> and <http://psh.ntkcz.cz/skos/PSH10340> are connected by skos:related and skos:broaderTransitive! <i>Checks if concepts are connected by clashing semantic SKOS relations:</i> <ul style="list-style-type: none">• connected by skos:related and skos:broaderTransitive• connected by skos:related and skos:narrower



Vizualizace hesla PSH (software PoolParty)

Prohlížení PSH



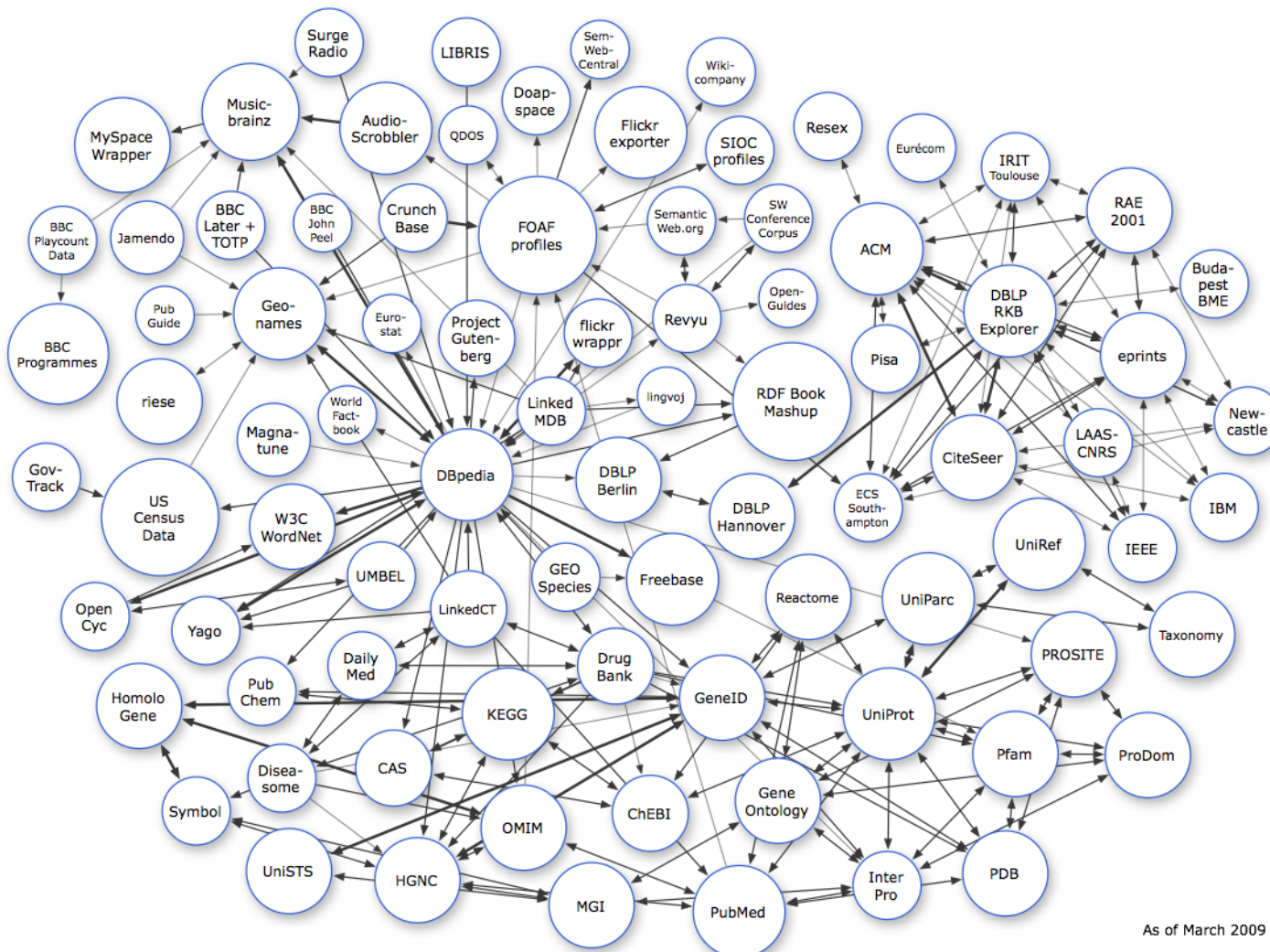
NTK

50°6'14.083"N, 14°23'26.365"E

Národní technická knihovna
National Technical Library



210 mm



As of March 2009

zdroj: http://www4.wiwiw.fu-berlin.de/bizer/pub/lod-datasets_2009-02-27.png

NTK

50°6'14.083"N, 14°23'26.365"E
Národní technická knihovna
National Technical Library



210 mm

Děkujeme za pozornost.

psh@techlib.cz