



národní  
úložiště  
šedé  
literatury

## **Indexační pravidla pro práci s Polytematickým strukturovaným heslářem (PSH)**

Skolková, Linda  
2007

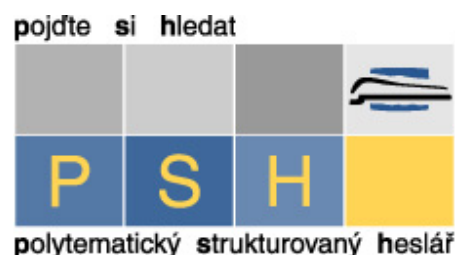
Dostupný z <http://www.nusl.cz/ntk/nusl-170822>

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

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

Datum stažení: 17.07.2024

Další dokumenty můžete najít prostřednictvím vyhledávacího rozhraní [nusl.cz](http://nusl.cz) .



## Indexační pravidla pro práci s Polytematickým strukturovaným heslářem (PSH)

*Poznámka:* Tento dokument vychází z návrhů obsažených v diplomové práci věnované PSH, která byla obhájena na Ústavu informačních studií a knihovnictví Filozofické fakulty Univerzity Karlovy v Praze v květnu 2007:

SKOLKOVÁ, Linda. *Polytematický strukturovaný heslář (Polythematic Structured Subject Heading System)*. Praha, 2007. iv, 172 s., 27 s. příl. + 1 CD-ROM. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí diplomové práce Ing. Miloslav Nič, PhD. K dispozici také na World Wide Web: <[http://www.skolkova.net/diplomka/SKOLKOVA\\_Linda\\_DP\\_PSH.pdf](http://www.skolkova.net/diplomka/SKOLKOVA_Linda_DP_PSH.pdf)>.

### *Předmluva*

Tato indexační pravidla představují základní metodický materiál pro indexátory pracující s Polytematickým strukturovaným heslářem (dále PSH či heslář).

Podnětem k vytvoření těchto pravidel se staly požadavky indexátorů, zejména těch, kteří s heslářem teprve začínají pracovat a využívat jej v rutinní indexátorské praxi.

Cílem explicitní formulace indexačních pravidel je zabezpečení konzistentní indexace hesly PSH v konkrétních institucích, v nichž je heslář využíván.

Základní struktura indexačních pravidel vychází ze struktury obdobných indexačních pravidel publikovaných v českém prostředí, zejména z [*Český teologický tezáur*], doplnkově pak z [*Český pedagogický tezáur*] a [*Tezáur ALIMIS*]. Vzhledem k tomu, že Polytematický strukturovaný heslář vykazuje řadu znaků typických pro tezaury, byla při vytváření tohoto návrhu zvolena jako východisko právě indexační pravidla pro oblast tezaurů, které lze zařadit mezi deskriptorové selekční jazyky (tj. předmětové selekční jazyky deskriptorového typu, viz [SCHWARZ, 1999, s. 11]).

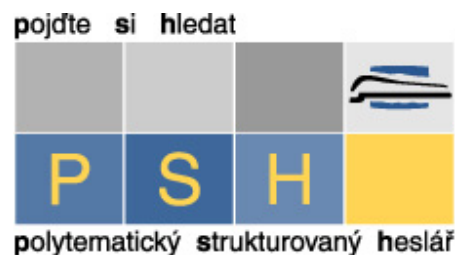
Indexační pravidla jsou zaměřena na českou část Polytematického strukturovaného hesláře. S anglickou částí hesláře lze pracovat obdobným způsobem.

Tento dokument je rozčleněn do tří základních částí: *Úvod*, *Indexace dokumentů hesly PSH* a *Aktualizace PSH*.

Konkrétní příklady uváděné v těchto indexačních pravidlech vycházejí z Polytematického strukturovaného hesláře verze 2.1.

### *1. Úvod*

Cílem této části textu je stručně objasnit problematiku věcného zpracování informací s důrazem na předmětové pořádání informací. Nejprve je podáno vysvětlení použitých termínů. Dále je pozornost věnována problematice obsahové analýzy, identifikace pojmů



a výběru hesel. Do uvedeného kontextu je také zařazen Polytematický strukturovaný heslář. V poslední části úvodu jsou zahrnuty vybrané položky literatury doporučené k dalšímu studiu.

### 1.1 Terminologické poznámky

Za **dokument** je považován „*informační pramen tvořený nosičem informací a množinou informací na něm fixovaných a sloužící k přenosu dat v čase a prostoru*“ [TDKIV, termín *dokument*]. Za podstatnou charakteristiku jsou přitom pokládány zaznamenané informace, konkrétní typ či druh nosiče informací je (zejména s ohledem na rozvoj elektronického publikování, a to zejména v prostředí internetu) pokládán na druhotný.

**Věcným zpracováním (pořádáním) informací** rozumíme „*specifický proces pořádání informací, při kterém jsou informace získané obsahovou analýzou dokumentu vyjádřeny pořádacími znaky věcných selekčních jazyků*“ [TDKIV, termín *věcné pořádání informací*].

**Předmětovým pořádáním informací** je „*druh věcného pořádání informací, při kterém je obsah dokumentu vyjádřen pomocí abecedně uspořádaných hesel, popř. lexikálních jednotek*“ [TDKIV, termín *předmětové pořádání informací*].

**Obsahovou analýzou** je míněna „*analýza obsahu dokumentu zahrnující metody a pravidla pro stanovení tematiky dokumentu, příp. časového a prostorového hlediska, čtenářského určení a formy dokumentu*“ [TDKIV, termín *obsahová analýza*].

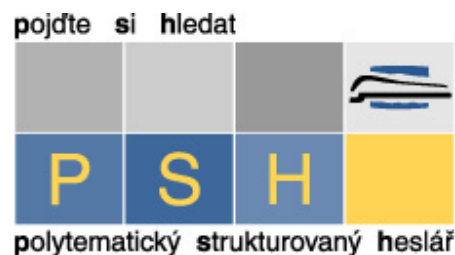
**Indexace** znamená „*proces vyjádření obsahu dokumentu pomocí prvků selekčního jazyka, obvykle s cílem umožnit zpětné vyhledávání*“ [TDKIV, heslo *indexace*].<sup>1</sup>

Termínem **pojem** označujeme „*myslenkový útvar odrážející podstatné vlastnosti předmětů, jeví objektivní skutečnosti*“ [SKOLEK, 2007].

**Lexikální jednotkou** máme na mysli „*sekvenci písmen, číslic a dalších specifických znaků používaných pro označení určitého pojmu*“ [TDKIV, termín *lexikální jednotka*]. Vzhledem k tomu, že lexikální jednotka je vždy vztažena k určitému lexiku (tj. souboru slovní zásoby určitého jazyka), je v tomto kontextu z praktických důvodů vhodnější užívat výraz **termín** [SKOLEK, 2007].

**Preferovaným termínem** je „*lexikální jednotka předmětového selekčního jazyka užívaná závazně při indexaci k reprezentaci určitého pojmu v dokumentu nebo rešeršním požadavku*“ [TDKIV, termín *preferovaný termín*]. V případě Polytematického strukturovaného hesláře je preferovaný termín ekvivalentní **heslu**, tedy nejkratšímu jazykovému vyjádření věcného obsahu práce [SKOLEK, 2007].

<sup>1</sup> Ekvivalentně lze využít termín *indexování*, popř. *heslování*. Termín *heslování* upřednostňoval např. J. TOMAN [1966].



**Nepreferovaný termínem** je „*ekvivalent nebo kvaziekvivalent preferovaného termínu; nepreferovaný termín není dokumentu přiřazován, ale slouží jako uživatelský vstup do řízeného slovníku odkazující prostřednictvím vztahu ekvivalence k příslušnému preferovanému termínu*“ [TDKIV, termín nepreferovaný termín].

V případě potřeby objasnit další v textu použité termíny doporučujeme využít *České terminologické databáze z oblasti knihovnictví a informační vědy* [TDKIV].

## 1.2 Věcné zpracování dokumentů

Věcné zpracování dokumentů je složeno ze tří základních kroků, které se však často vzájemně prolínají. Jedná se o obsahovou analýzu, identifikaci pojmů a termínů a výběr hesel<sup>2</sup> (volně podle [Český teologický tezaurus] a [ČSN ISO 5963]).

### 1.2.1 Obsahová analýza

Důkladně provedená obsahová analýza představuje základ pro kvalitní indexaci, která je nezbytná pro kvalitní vyhledávání v daném informačním systému. Při obsahové analýze je podle [ČSN ISO 5963] doporučeno vycházet zejména z následujících částí dokumentů (na příkladě knižních dokumentů):

- název<sup>3</sup>;
- obsah<sup>4</sup>;
- referát, byl-li zpracován;
- úvod, úvodní věty kapitol nebo odstavců a závěr;
- slova nebo skupiny slov, které jsou podtrženy nebo zvýrazněny např. neobvyklým typem písma.

Obsahová analýza má podstatný vliv na kvalitu indexace a následně zpětného vyhledávání. Její nesprávné provedení může vést k chybám dvojího druhu:

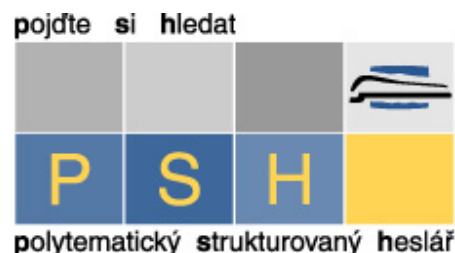
- ztráta informací;
- indexace nerelevantními termíny.

---

<sup>2</sup> Výraz *výběr hesel* se již přímo vztahuje k PSH, v jiných případech (např. u tezurů) se může jednat o výběr *deskriptorů*.

<sup>3</sup> Název však může být zavádějící, není proto vhodné při indexaci vycházet pouze z názvu.

<sup>4</sup> Ve smyslu anglického výrazu *table of contents*.



## 1.2.2 Identifikace pojmů a termínů

Na základě výsledků obsahové analýzy dokumentu indexátor identifikuje nejdůležitější pojmy a termíny, které nejvýstižněji charakterizují tematiku indexovaného dokumentu. Při této činnosti lze vycházet z kontrolních otázek, které uvádí příslušná norma [ČSN ISO 5963] či příručky vztahující se k problematice obsahové analýzy a indexace (viz kap. 1.4), popř. doplnit další otázky, o nichž je známo nebo lze předpokládat, že jsou pro danou instituci a uživatele jejího informačního systému důležité.

Míra úplnosti indexace je závislá na účelu daného informačního systému, v němž je selekční jazyk využit. Indexátor by se vždy měl řídit zásadou nejužšího pojmu, tj. vybrat takové termíny, které nej přesněji vyjadřují tematiku dokumentu.<sup>5</sup> Termíny vyjadřující obecnější pojmy lze záznamům dokumentů přiřazovat pouze ve výjimečných případech (za předpokladu nepříznivého vlivu na vyhledávání či je-li specifitější pojem v dokumentu zmíněn pouze okrajově).

Při obsahové analýze by se indexátor měl seznámit s nejdůležitějšími termíny, které jsou použity v dokumentu, a s jeho tématem v dostupných referenčních zdrojích. K vyhledání termínů mohou být využity např. odborné slovníky, u neologismů také internetové vyhledávače (zejména k ověření skutečně používaných variant termínů dle četnosti výskytu). Přednost je dávána termínům, které jsou preferovány odbornou veřejností.

## 1.2.3 Výběr hesel

Pro výběr hesel, tj. termínů konkrétního předmětového selekčního jazyka, by měl být indexátor obeznámen s jeho celkovou organizací a obsahem. Zároveň by měl při své práci vycházet z indexační politiky instituce provozující daný informační systém. Tato politika by měla být explicitně vyjádřena, nejlépe v textové podobě. Může tvořit dílčí součást politiky věcného zpracování fondu, resp. politiky zpracování fondu.

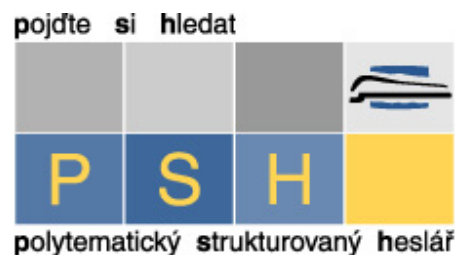
Hesla jsou volena na základě výsledků obsahové analýzy a identifikace pojmů a termínů tak, aby byl daný dokument co nejvýstižněji popsán z hlediska své tematiky.

## 1.3 Heslář

### 1.3.1 Základní charakteristika

Polytematický strukturovaný heslář je česko-anglický řízený a měnitelný slovník lexikálních jednotek ze všech základních oblastí lidského poznání. Slouží k vyjádření věcného obsahu dokumentů a ke zpětnému vyhledání dokumentů na základě věcných kritérií.

<sup>5</sup> R. FUGMANN [1993, s. XI, 88-90] pro tento typ indexace používá termín *mandatorní indexace*.



Heslář byl vytvořen knihovníky ve spolupráci s odborníky z jednotlivých oborů, které pokrývá. Ve verzi 2.1 je heslář rozčleněn do celkem 44 tematických řad a obsahuje přes 13 000 hesel.

Správci hesláře doporučují jej používat zároveň s dalšími selekčními jazyky, ať již řízenými (např. Mezinárodní desetinné třídění) či neřízenými (např. volně tvořená klíčová slova), aby byl uživateli zajištěn nejvyšší možný komfort při věcném vyhledávání informací.

V hesláři se vyskytují tři základní druhy apriorních pojmových vztahů – vztah hierarchie, vztah ekvivalence a vztah asociace.

Vztahy hierarchie existují mezi hesly vyjadřujícími užší a širší pojmy, ekvivalenční vztahy pak mezi hesly a jejich synonymy, popř. kvazisynomymy. V řadě případů je kvůli přehlednosti hesláře jako celku využito hierarchizace vztahu ekvivalence, tj. některé termíny vyjadřující užší pojmy jsou zařazeny jako termíny nepreferované u hesel vyjadřujících odpovídající širší pojmy. Vztahy asociace jsou využity v případech, kdy je pokládáno za přínosné indexátory a/nebo uživatele upozornit na další příbuzná hesla, která jsou umístěna v jiné části hesláře.

Nová hesla do hesláře přidávají jeho správci na základě návrhů indexátorů, koncových uživatelů a externích spolupracovníků (specialistů v jednotlivých oborech). Jednotlivé instituce používající heslář mohou do hesláře přidávat vlastní hesla (např. podrobněji zachytit oblasti, které jsou v hesláři zastoupeny pouze jedním či několika hesly). Zajištění možnosti přechodu daného informačního systému na novou verzi hesláře či průběžných aktualizací je však v těchto případech v kompetenci jednotlivých institucí, správce hesláře v tomto procesu může mít pouze poradenskou a konzultační roli.

Heslář lze zařadit mezi postkoordinované systémy. Znamená to, že složené pojmy jsou vyjádřeny skládáním jednotlivých hesel (na rozdíl např. od klasických systémů předmětových hesel, které jsou zařazovány mezi prekoordinované systémy).

### 1.3.2 Termíny

Význam výrazu *termín* v oblasti selekčních jazyků a věcného vyhledávání informací obecně je vyložen v kap. 1.1 Terminologické poznámky. V konkrétním případě Polytematického strukturovaného hesláře sem patří všechny termíny obsažené v hesláři, tj. jak preferované (hesla), tak nepreferované termíny (odkazy typu viz).

Preferované termíny (hesla) jsou výrazy, které se používají pro indexaci. Jedná se o jednoslovné (např. *historie bi*) nebo víceslovné neboli složené (např. *informační prameny in*)

pojďte si hledat



polytematický strukturovaný heslář

termíny, které jsou doplněny dvouznakovými zkratkami, které označují příslušnost hesla do tematické řady.

Nepreferované termíny (odkazy typu viz) jsou výrazy, které se nepoužívají pro indexaci. Indexátorům slouží k vyhledání preferovaných termínů použitelných pro indexaci, a uživatelům pro účely vyhledávání. V hesláři se nacházejí termíny, které vyjadřují pojmy konkrétní i abstraktní povahy (v následující části textu je využito rozčlenění dle [ČSN 01 0193]):

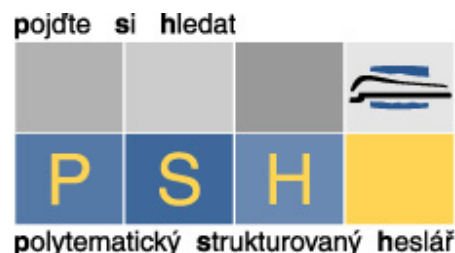
### Konkrétní entity:

- předměty, věci a jejich fyzické části:
  - *lesnické stroje ze*
  - *balkony au*
- materiály:
  - *sklo sp*
  - *dřevo ze*

### Abstraktní entity:

- činnosti a události:
  - *odstřeďování ch*
  - *odkalování ts*
- abstraktní entity a vlastnosti věcí, materiálů nebo činností
  - *pevnost fy*
  - *hmotnost fy*
- obory nebo vědní disciplíny
  - *fyzika fy*
  - *chemie ch*
- měřicí jednotky [v PSH nejsou zařazeny konkrétní měřicí jednotky, ale pouze obecná označení tříd měřicích jednotek]





- fyzikální jednotky *fj*

## Individuální entity neboli “jednotkové třídy” vyjádřené vlastními jmény

- *Galaxie (Mléčná dráha) as*

Hesla jsou tvořena substantivy, která mohou (např. *mezigalaktická hmota as*), ale nemusí (např. *mlhoviny as*) být rozvíta adjektivy. Jsou-li použita adjektiva, jsou přednostně uváděna před substantivy (viz příklady výše). V ojedinělých případech následují adjektiva až za substantivy. Jedná se mj. o odbornou chemickou terminologii (např. *oxid ubličitý ch*).

V některých případech jsou použita substantiva s předložkovou vazbou (např. *čas v ekonomii ev*), popř. kombinace adjektivní a předložkové vazby (např. *výpočetní metody ve fyzice fj*). Vyskytují se rovněž hesla, v nichž je substantivum rozvíta dvěma adjektivy, z nichž jedno se nachází před substantivem a druhé za substantivem (např. *jazykové obory základní ja*). V případech, kdy byl v předchozích verzích použit odlišný tvar hesla, jsou tato hesla z důvodu zpětné kompatibility do hesláře zařazena jako nepreferované termíny (např. *měření teploty a měření teplot viz teplota fj*).

Príslovce se v rámci hesel vyskytují ojediněle (např. *rozvodny velmi vysokého napětí en*). Slovesa nejsou do hesláře zařazována vůbec.

Zkratky a akronymy<sup>6</sup> se v hesláři vyskytují v pozici samostatných hesel (např. *DNA ch*) nebo jako nepreferované termíny (např. nepreferovaný termín *ASŘ* odkazující na heslo *automatizované systémy řízení ev*).

Pokud se týká užití singuláru a plurálu, počítatelná substantiva se v zásadě vyjadřují plurálem (např. *grafy ob*), ale v některých případech je použit singulár (např. *výpověď ja*). U nepočítatelných substantiv je přednostně užit singulár (např. *pivo pp*). Pokud lze předpokládat, že uživatelé budou pokládat např. určitou látku za třídu, pak je použit plurál (např. *jedy ch*). Lexikální jednotky vyjadřující abstraktní pojmy jsou přednostně uvedeny v singuláru (např. *historie hì*), pouze výjimečně v plurálu (např. *chemické reakce ch*). V případě výrazné odlišnosti singuláru a plurálu je v konkrétních případech přidán nepreferovaný termín (např. *dítě viz děti so*).

V případech homografie či polysémie jsou v současnosti v hesláři řešeny tak, že dané (např. dva) pojmy jsou vyjádřeny dvěma různými lexikálními jednotkami (např. *krídlo bi*

<sup>6</sup> Akronym je „zkratkové vlastní jméno vzniklé spojením začátečních písmen několika slov (např. UNESCO)“ [Akademický slovník, s. 33]. Zkratka je „ustálené zkrácení slova nebo slov“ [Slovník spisovného, s. 760].



pojdte si hledat



polytematický strukturovaný heslář

a *keřídla letadla sr*). Perspektivně počítají správci hesláře se zavedením relátorů (kvalifikátorů, závorkového výkladu), které se nyní vyskytují v hesláři pouze ojediněle (např. *strukturalismus (filozofie) fi* a *strukturalismus (jazykověda) ja*).

V hesláři jsou používány nejrozšířenější (eventuálně kodifikované) pravopisné tvary lexikálních jednotek. Méně rozšířené termíny (včetně zastaralých), popř. slangové výrazy či žargon, jsou uváděny jako nepreferované termíny (např. *keřylčičníky viz oxidy ch*).

Pokud je pro daný pojem používán běžný název, je přednostně použit běžný název (např. *jaderná chemie ch*, nepreferovaný termín *nukleární chemie*). Pokud je však v odborné terminologii upřednostňován vědecký název, je i v hesláři použit jako heslo název vědecký (např. *substituce ch*, nepreferovaný termín *záměna*). Pokud je pro daný pojem užíván pouze vědecký název, je použit tento název (např. *endoplazmatické retikulum bi*).

Formální termíny nejsou součástí hesláře. Hesla typu *mapy gr* či *učebnice pe* je doporučeno používat pouze v případech, kdy jimi lze vyjádřit tematiku dokumentu, nikoliv jeho formu.<sup>7</sup>

### 1.3.3 Struktura

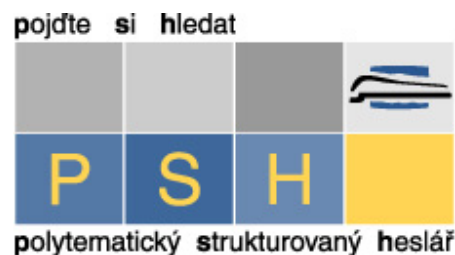
Polytematický strukturovaný heslář obsahuje tři základní struktury – hierarchickou, asociační a ekvivalenční. Díky tomu se lze v hesláři snadno orientovat a tím je usnadněna indexace i vyhledávání.

Pro hierarchickou strukturu je charakteristický vztah nadřazenosti a podřazenosti. Ve verzi 2.1 hesláře se na nejvyšších místech hierarchické struktury nachází celkem 44 hesel označujících tematické řady. U hesel na všech hierarchických úrovních jsou za příslušným termínem uvedeny dvouznakové zkratky, které označují příslušnost hesla do dané tematické řady. Např. heslo *termika fj* náleží do tematické řady *fyzika fj*, tj. je podřazena heslu *fyzika fj*. Bezprostředně nadřazeným heslem však není heslo *fyzika fj*, ale heslo *termodynamika fj*. Díky dvouznakovým zkratkám tak lze i u hesel na nižších hierarchických úrovních snadno zjistit příslušnost k dané tematické řadě.

Asociační struktura naznačuje příbuznost hesla vůči jinému heslu, které není v hesláři uvedeno v pozici nepreferovaného termínu ani není vůči danému heslu v hierarchicky nadřazeném nebo podřazeném postavení (např. *balneologie zd* viz též *minerální prameny vo* a *balneoterapie zd*).

Ekvivalenční struktura představuje vazbu mezi heslem a nepreferovaným termínem (např. *péče o kůži viz péče o pleť zd*).

<sup>7</sup> Např. heslem *učebnice pe* může být popsán dokument pojednávající o tvorbě učebnic matematiky.



V případě všech výše uvedených struktur se jedná o oboustranné (reciproční) vazby.

V hesláři se vyskytují dva základní druhy vztahů hierarchie:

- generický vztah (např. *savci bi - chobotnatci bi*)
- partitivní vztah (např. *dýchací systém bi – plíce bi* či *psychologické směry ps – analytická psychologie ps*)

Vzhledem k tomu, že v hesláři se nevyskytují identifikátory, nevyskytují se v něm ani kauzální vztahy v hierarchii, tj. případy, kdy je vyjádřen vztah mezi obecným substantivem (třídou) a individuálním případem (instancí) označeným vlastním jménem.

Heslář je uspořádán do stromové struktury. Z ní lze odvozením získat např. abecední seznam hesel, nepreferovaných či asociovaných termínů, v závislosti na využívaném programovém vybavení lze v hesláři také vyhledávat podle řetězců a podřetězců.

#### 1.4 Doporučená literatura k dalšímu studiu

Podrobnější informace k tématu věcného pořádkání informací s důrazem na předmětové selekční jazyky lze nalézt např. v následujících zdrojích:

*The American Society of Indexers* [online]. Wheat Ridge (CO, USA) : American Society of Indexers [cit. 2007-03-30]. Dostupné na World Wide Web: <<http://www.asindexing.org/>>.

DRTINA, Jaroslav. *Předmětový katalog*. Praha : Státní pedagogické nakladatelství, 1957. 150 s.

FUGMANN, Robert. *Subject analysis and indexing : theoretical foundation and practical advice*. Frankfurt am Main : Indeks Verlag, 1993. 250 s. Textbooks for Knowledge Organization, vol. 1. ISBN 3-88672-500-6.

HYHLÍKOVÁ, Věra. *Informační analýza dokumentu*. Praha : ÚVTEI, 1984. 80 s. Učební texty, sv. 19.

*ISKO : International Society for Knowledge Organization* [online]. [S.l.] : cISKO 2004-2007, last updated 2007.02.26 [cit. 2007-03-30]. Dostupné na World Wide Web: <<http://www.isko.org/>>.

KOVÁŘ, Blahoslav. *Obsahová analýza dokumentu*. Praha : ÚVTEI, 1974. 44 s. Metodické letáky, sv. 105.

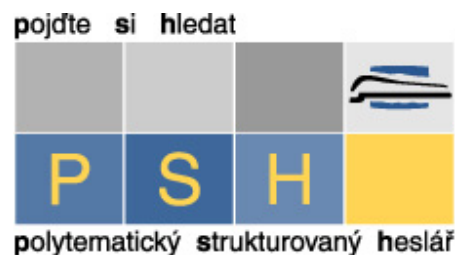
KOVÁŘ, Blahoslav. *Základy obecné metodiky věcného pořádkání informací*. Praha : ÚVTEI, 1978. 65 s. Metodický leták, sv. 130.

OLSON, Hope A.; BOLL, John J. *Subject analysis in online catalogs*. 2<sup>nd</sup> ed. Englewood (CO, USA) : Libraries Unlimited, 2001. xv, 33 s. ISBN 1-56308-800-2.

ŠAUPERL, Alenka. *Subject determination during the cataloging process*. Lanham (MD, USA); London (Velká Británie) : Scarecrow Press, 2002. vii, 173 s. ISBN 0-8108-4289-0.

## 2. Indexace dokumentů hesly PSH

Dokumenty nemusejí být popsány pouze jedním heslem, ve většině případů je vhodné použít více hesel. Důvodem je snaha zachytit co nejlépe tematiku popisovaného



dokumentu. Indexace dokumentů hesly PSH by se vždy měla vycházet z potřeb uživatelů informačního systému, v němž jsou hesla využita.

## 2.1 Výběr konkrétních hesel

K indexaci dokumentů se používají pouze hesla (preferované termíny). Jednotlivá hesla jsou uspořádána hierarchicky, tj. do stromové struktury. Každé heslo je zařazeno pouze na jedno místo v hierarchii; je uplatněn princip tzv. *monohierarchie*. Každé heslo je také doplněno dvouznakovou zkratkou, která označuje jeho příslušnost do tzv. *řady*. Jako *řadu* označujeme skupinu nejobecnějších hesel, která se podrobněji člení.

Při výběru konkrétních hesel je nezbytné brát v úvahu kontext, v němž se vyskytují – kromě nadřazených a podřazených hesel také hesla asociovaná a teprve na základě vyhodnocení příslušného kontextu vybrat heslo či hesla vhodně reprezentující pojem obsažený v indexovaném dokumentu.

Z jedné hierarchické linie by mělo vždy, je-li to možné a účelné, být vybráno pouze jedno heslo.

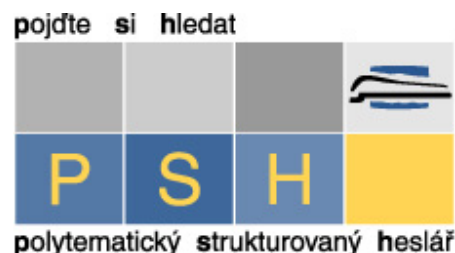
Významově širší heslo volíme v těchto případech:

- dokument by bylo možné popsat všemi hesly z nižší hierarchické úrovně (např. dokument, který se týká listnatých stromů, jehličnatých stromů a okrasných dřevin, bude indexován heslem *dřeviny ze*);
- významově užší heslo není v hesláři obsaženo (např. dokument pojednávající o unitermech bude popsán pomocí hesla *selekční jazyky deskriptorové in*).

Významově užšímu heslu pak dáváme přednost v případech, kdy:

- je dokument převážně věnován užšímu tématu, širší téma je zmíněno pouze okrajově (např. dokument pojednávající o převážně o superpočítačích a pouze okrajově o počítačích obecně bude popsán heslem *superpočítače vt* nebo dokument, který se podle názvu týká matematiky obecně, z obsahu je však zřejmé, že pojednává pouze o diferenciálních rovnicích, bude popsán heslem *diferenciální rovnice ma*);
- dokument by bylo možné popsat pouze některými hesly z nižší hierarchické úrovně (např. dokument popisující problematiku mechanického a pneumatického třídění surovin, avšak nikoliv již třídění hydraulického, bude vyjádřen hesly *mechanické třídění ts* a *pneumatické třídění ts*, nikoliv *třídění surovin ts*).

V případě, že konkrétní heslo není v PSH obsaženo, je nejprve potřeba se pokusit pojem vyjádřit jiným způsobem – např. rozložit jej na sémantické složky, k nimž lze v hesláři nalézt odpovídající verbální vyjádření, nebo dát přednost heslu na vyšší úrovni obecnosti.



Pokud tyto varianty řešení nelze aplikovat, může být daný výraz navržen jako kandidát na nové heslo.

### 2.1.1 Určení významnosti hesel

Doporučeným pořadím jednotlivých hesel v záznamu dokumentu je pořadí významnosti, tj. hesla, která reprezentují nejdůležitější pojmy obsažené v dokumentu, jsou uváděna jako první, hesla reprezentující méně důležité pojmy pak na dalších místech.

V závislosti na možnostech a účelu konkrétního informačního systému je také možné rozlišovat významnost jednotlivých hesel vzhledem k obsahu dokumentu explicitně, např. formou rozlišení hlavních a vedlejších hesel, resp. témat (tedy zavést jednoduchý způsob vážení<sup>8</sup>). V těchto případech lze doporučit umožnit uživatelům vyhledávání podle hlavních a vedlejších hesel a rovněž toto rozčlenění vhodně zobrazit v jednotlivých záznamech (např. v podobě popisek *hlavní téma* a *vedlejší téma*).

### 2.1.2 Počet hesel

Počet hesel přidělený jednomu dokumentu není omezen. Je-li to možné, měla by však být zachována zásada proporcionality, tj. rozdíly mezi počty hesel přidělených různým dokumentům by neměly být výrazné. Doporučené počty přidělovaných hesel lze rovněž stanovit v indexační politice.

Není-li v dané instituci explicitně stanovena indexační politika, lze za optimální počet hesel přidělených na jeden dokument obecně považovat jedno až šest hesel v závislosti na tematice konkrétního dokumentu a na předpokládaných potřebách uživatelů konkrétního informačního systému, v němž je PSH využíván k vyhledávání. Výše uvedený počet vychází z praxe indexátorů ve Státní technické knihovně.

### 2.1.3 Zjištění konkrétních hesel

Ke zjištění konkrétních hesel vhodných k zachycení věcného obsahu dokumentu lze využít jednak stromové struktury hesláře, která umožňuje snadnou identifikaci nadřazených a podřazených hesel, jednak vztahů ekvivalence (odkazů z nepreferovaných termínů na hesla, tj. termíny preferované) a asociace (odkazů na tematicky příbuzná hesla). V hesláři je možné také jednotlivé termíny pomocí dotazů vyhledat a následně si je nechat zobrazit ve stromové struktuře, aby byl zřejmý kontext vyhledaného hesla.

---

<sup>8</sup> Návrhy systémů vah včetně stručného hodnocení jsou k dispozici např. v [AITCHINSON; GILCHRIST; BAWDEN, 1997, s. 89-90]. Se systémem vážení aplikovaným v oblasti věcného vyhledávání dokumentů se lze setkat např. v databázi MEDLINE, kde jsou hesla ze systému *Medical Subject Headings*, tj. MeSH, označována jako *major* a *minor concepts*, tedy hlavní a vedlejší pojmy.

pojdte si hledat



polytematický strukturovaný heslář

#### 2.1.4 Prekoordinace a postkoordinace

Prekoordinace je uplatňována v případech, kdy je pojem vyjádřen složeným termínem. V případě, že pro daný pojem existuje v hesláři odpovídající složený termín, je dána při indexaci přednost tomuto termínu. V případě, že pro daný pojem v hesláři odpovídající složený termín neexistuje, je uplatněn princip postkoordinace, tj. dokumentu jsou přidělena dvě hesla (popř. více hesel), která v souhrnu vyjadřují daný pojem. Vzhledem ke struktuře hesláře, která se blíží spíše tezauru než předmětovým heslům, je silně upřednostňován princip postkoordinace. Podmínkou jeho využití je předpoklad, že při užití kombinace hesel nedojde ke ztrátě původního významu. Princip postkoordinace lze úspěšně využít v případě komplexních témat.

Příklady uplatnění principu prekoordinace:

- *stavební fyzika st*, nikoliv *stavebnictví st* + *fyzika fy*
- *analytická chemie ch*, nikoliv *analýza ve* + *chemie ch*

Příklad uplatnění principu postkoordinace:

- téma *chemická informatika* vyjádřené hesly *chemie ch* a *informační věda in*

#### 2.1.5 Indexační chyby

Při praktické indexaci může docházet k chybám, které lze rozdělit na dva základní typy – věcné chyby a formální chyby. Mezi věcné chyby patří např. nesprávné stanovení významu a rozsahu hesla, využití příliš obecného nebo naopak příliš specifického hesla, záměna formy a obsahu dokumentu, indexace okrajových témat či nesprávné provedení obsahové analýzy. Mezi formální chyby řadíme např. použití nepreferovaného termínu místo hesla.<sup>9</sup>

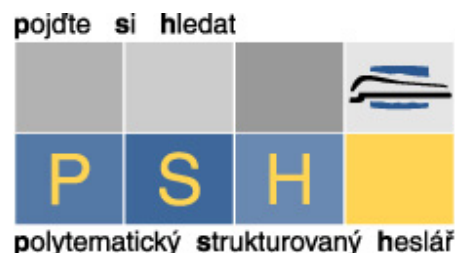
Svou roli hraje také subjektivita indexátora, kterou ovlivňuje řada faktorů (mj. znalosti a zkušenosti indexátora, podmínky práce apod.).

### 3. Aktualizace PSH

Vzhledem k tomu, že řada oborů se dynamicky vyvíjí, je heslář pravidelně aktualizován svými správci, tj. pracovníky Státní technické knihovny. Návrhy na doplnění hesel či úpravu stávajících hesel jsou velmi vítány. K zaslání návrhů je možno použít elektronickou poštu ([psh@stk.cz](mailto:psh@stk.cz)), klasickou poštu (Státní technická knihovna, oddělení

---

<sup>9</sup> Tento typ formální chyby však nenastává v automatizovaných systémech, v nichž je indexátorovi povolen pouze výběr hesla, nikoliv nepreferovaného termínu. Tomuto typu chyby lze zamezit softwarovými prostředky.

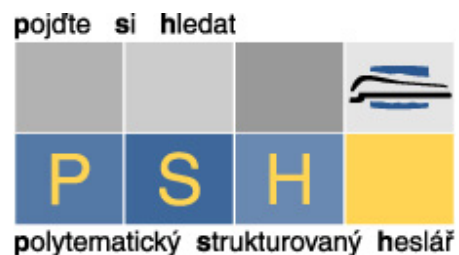


speciálních služeb a PSH, Mariánské nám. 5, pošt. přihrádka 206, 110 01 Praha 1) nebo telefon (+420 221 663 465). Základní informace o hesláři včetně všech kontaktních údajů jsou také k dispozici v rámci webové prezentace hesláře, která je dostupná na adrese <http://www.stk.cz/psh.html>.

### **Použitá literatura**

- AITCHINSON, Jean; GILCHRIST, Alan; BAWDEN, David. 1997. *Thesaurus construction and use : a practical manual*. 3<sup>rd</sup> ed. London : Aslib, c1997. xvi, 211 s. ISBN 0-85142-390-6.
- Akademický slovník cizích slov*. Kolektiv autorů pod vedením Věry Petráčkové a Jiřího Krause. Praha : Academia, 2000. ISBN 80-200-0607-9.
- Český pedagogický tezáurus*. Zpracovali Naděžda Baránková, Antonín Dušek, Jiří Haškovec. 2. upravené vyd. Praha : Ústav pro informace ve vzdělávání, 2002. 327 s. ISBN 80-211-0435-X.
- Český teologický tezarus : verze 1.0 : indexační pravidla : verze 1.2*. Šest. Adéla Baková. Vyd. 1. Praha : Evangelická teologická fakulta Univerzity Karlovy, 1998. 28 s., příl.
- ČSN 01 0193. *Pokyny pro vypracování a rozvíjení jednojazyčných tezarurů*. Praha : Český normalizační institut, 1995. 49 s.
- ČSN ISO 5963 (01 0174). *Dokumentace. Metody analýzy dokumentů, určování jejich obsahu a výběru lexikálních jednotek selekčního jazyka*. Zpracovatel Karel Pech. Praha : Český normalizační institut, 1995.
- FUGMANN, Robert. 1993. *Subject analysis and indexing : theoretical foundation and practical advice*. Frankfurt am Main : Indeks Verlag, 1993. 250 s. Textbooks for Knowledge Organization, vol. 1. ISBN 3-88672-500-6.
- SCHWARZ, Josef. 1999. *Vývoj teorie a praxe tezarurů v České republice : nástin dějin deskriptorových selekčních jazyků v bývalém Československu se zaměřením na vývoj teoretických, metodických a normativních aspektů tvorby tezarurů*. ix, 121 s. Praha : vlastním nákladem, 1999. Diplomová práce. Univerzita Karlova v Praze, Filozofická fakulta, Ústav informačních studií a knihovnictví. Vedoucí diplomové práce Peter Pálka.
- SKOLEK, Jaroslav. 2007. *Osobní komunikace*. 15. 3. 2007.
- Slovník spisovného jazyka českého*. IV. V-Ž, Doplnky a opravy. Praha : Academia, 1971. 1007 s.





*TDKIV* : Česká terminologická databáze z oblasti knihovnictví a informační vědy [online databáze]. 2003- . Praha : Knihovnický institut NK ČR, 2003- [cit. 2007-01-20]. Databáze vznikla v letech 2001-2002 v rámci projektu podpořeného grantem MK ČR. Dostupná na World Wide Web: <<http://sigma.nkp.cz/cze/ktd>>.

*Tezaurus ALIMIS*. Praha : Výzkumný ústav potravinářského průmyslu; Středisko technických informační potravinářského průmyslu, 1990.

TOMAN, Jiří. 1966. Principy a systémy heslování a třídění : (úvod do pořádání informací). Praha : Státní technická knihovna, 1966. 187 s.