

Věcná katalogizace

PhDr. Milada Píšková

Věcná katalogizace

- Věcné pořádání informací (indexace)
 - dílčí proces vstupního zpracování informací, vychází z obsahové analýzy dokumentu t.j. informace jsou při něm pořádány podle obsahových údajů získaných informační analýzou

- **Obsahová analýza dokumentu**

- předpoklad kvalitního věcného zpracování
 - výsledek: obsahové údaje

- nejprve nutná intelektuální redukce nebo kondenzace textu

- dále převod lexikálních jednotek přirozeného jazyka do lexikálních jednotek jazyka informačního (selekčního jazyka)

- Vyžaduje: rozbor a pečlivou prohlídku těch míst dokumentu, které obsahují důležité obsahové údaje :

- název a podnázev
 - úvod , předmluva
 - obsah

- Prostředek věcného pořádní informací

- **selekční jazyk**

druh informačního jazyka, určený k formulaci identifikačních i obsahových údajů o dokumentech, k uložení dokumentu a k formulaci dotazu uživatele na výstupu z informačního systému

- selekční jazyky

Systematický selekční jazyk

- Systematický selekční jazyk je výrazový a vyjadřovací prostředek systematického pořádní informací

Obsah dokumentu lze vyjádřit prostřednictvím kódovaných znaků umělého selekčního jazyka.

Tzn.že získané poznatky , lze zařadit do předem systematicky uspořádané soustavy lidského poznání

Příklad: Mezinárodní desetinné třídění(MDT)

Deweyho desetinné třídění (DDC)

Předmětový selekční jazyk

- Je výrazový a vyjadřovací prostředek předmětového pořádání informací
- Selekční obraz dokumentu je vyjádřen prostřednictvím lexikálních jednotek speciálně upraveného tj. formalizovaného přirozeného jazyka
- Výsledná formulace – předmětové heslo, případně skupina klíčových slov- je zařazena do seznamu těchto selekčních prvků

Prekoordinované selekční jazyky

- Se vyznačují slovníkem , který se skládá nejen z jednotlivých slov, ale ze sousloví , popřípadě vět. V souslovích a větách těchto selekčních jazyků jsou jednotlivá slova uspořádána již v průběhu indexování dokumentu, před uložením
- Lexikální jednotky získané obsahovou analýzou dokumentu jsou uváděny ve vztahu podle předem stanovených pravidel. Není možné do takto uspořádaného souboru prvků vstupovat , transformovat a využívat jeho jednotlivé části

- **Postkoordinované selekční jazyky**
- (slovo latinského původu postkoordinovanost znamená „přiřadit k sobě dodatečně“)

Používají při indexování dokumentu a také při jejich uložení vzájemně izolovaná slova.

Složené téma dokumentu je z nich skládáno přiřazováním teprve v průběhu vyhledávání příslušných dokumentů odpovídajících obsahu uživatelského dotazu

Současné selekční jazyky předmětových hesel lze charakterizovat jako jazyky prekoordinované se silným průnikem prvků postkoordinace

Univerzální selekční jazyky

umožňují pořádat , ukládat a vyhledávat informace z celého univerza lidského poznání a praxe

Příklad: Mezinárodní desetinné třídění(MDT)

Speciální selekční jazyky

umožňují pořádat , ukládat a vyhledávat z určitého vymezeného výseku univerza lidského poznání a praxe

Příklad: MASCH

Top- down

selekční jazyky konstruovány shora dolů,
od celistvé abstraktní představy ke konkrétní
implementaci

Bottom-up

selekční jazyky se slovníkem a vztahy
doplňovanými průběžně v procesu
používání
od jednotlivých lexikálních jednotek ke
kompletnímu slovníku

Historie systematických selekčních jazyků

- Souvisí s klasifikací věd
- Starověk

Aristoteles (4. st. př.n.l.)

Rozeznával -teoretické vědy(matematika,
fyzika, metafyzika)

- praktické vědy (etika,
ekonomika, politika)

- poetické vědy(poetika, rétorika,
umění)

Středověk – systém sedmi svobodných umění
(souvisí se školstvím)

- Počátek novověku
- Francis Bacon(1561-1626)

Klasifikuje vědy podle toho, kterými schopnostmi člověk poznává svět:

paměť (základ historie)

- přírodní dějiny
- dějiny lidstva

fantazie (základ poezie)

rozum (základ vědeckého poznání)

každý z oddílů je dále klasifikován

vliv na francouzské encyklopedisty

ovlivnil W.T.Harrise ,ten dále ovlivnil M.Deweyho(tvůrce

Desetinného třídění) – vliv na Mezinárodní desetinné třídění

Systematické pořádání informací

- nejstarší druh katalogu: systematické katalogy
- nejstarší pořádání informací vůbec

Nippur(kolem roku 2500 před n.l.)

knihovna uspořádána systematicky

Knihovna asyrského krále **Ašurbanipala**

(7.st.před n.l.)

Alexandrijská knihovna

Zenodotos **Kallimachos** (260-210 před n.l.)

katalog klasifikován na poezii, právo, filozofii, historii, rétoriku a díla různého obsahu

Středověké klášterní knihovny – systematické katalogy , knihy uspořádány na policích dle oborů

- Novověk

- Vynalezení knihtisku

 - nárůst literatury

 - růst knihoven

 - potřeba uspořádat literaturu

19. století – USA, Anglie

Rozvoj veřejných knihoven, potřeba sebevzdělávání

- významné systematické selekční jazyky

W.T.Harris

1870 – klasifikace pro knihovnu Veřejné školy v St.Louis

vyšel z Baconovy klasifikace věd

převrátil pořadí věd (na počátek položil vědu)

ovlivnil Deweyho a strukturu pozdějšího

Mezinárodního desetinného třídění

- Melvil **Dewey** (1851-1931)
- 1876- Deweyho klasifikace(DDT)
jeden z nejrozšířenějších systematických selekčních jazyků
dnes snaha nahradit jím MDT
přínos – zavedení číselného desetinného principu
zavedl systém pomocných znaků

všeobecnosti
filozofie
náboženství(teologie)
sociologie
filologie
přírodní vědy
užité vědy
umění
historie

- **Mezinárodní desetinné třídění (MDT)**
- Mezinárodní bibliografický ústav(r.1895 rozhodnuto o založení) Paul Otlet, Henri La Fontaine
 - úkol: vydávat bibliografické záznamy o světové tiskové produkci
 - první fáze přebírá Deweyho Desetinné třídění
 - dále se vytváří systém pomocných znaků
 - podrobnější rozpracování tabulek jednotlivých vědních oborů

r. 1905 – první úplné vydání MDT

• MDT se člení na deset základních tříd:

0 Všeobecnosti

1 Filozofie

2 Náboženství. Teologie

3 Společenské vědy

4 (neobsazena od r.1962)

5 Matematika. Přírodní vědy

6 Užité vědy.Lékařství. Technika

7 Umění

8 Jazykověda.Filologie.Literatury

9 Geografie. Biografické studie. Dějiny

- Každá tato hlavní třída stejně jako každá další hlavní třída MDT se člení připojováním dalších číslic na deset či méně oddílů nižšího řádu

Příklad:

6 Užití vědy. Lékařství. Technika

61 Lékařské vědy. Medicína

616 Patologie. Klinická medicína

616.8 Neurologie. Neuropatologie. Nervový
systém

616.85 Organické neurózy. Neuropatie

616.853 Epilepsie. Padoucnice

Pro přehlednost se píše za každou třetí číslicí tečka

- Příklad:

62 Technika

621.3 Elektrotechnika

621.38 Elektronika

621.383 Fotonky a fotoelektrické
články

621.383.292 Fotonásobiče

Hlavní (jednoduché) znaky MDT

znaky vzniklé postupným rozkladem znaku MDT

uvedeny v hlavních tabulkách

Složené znaky MDT

spojení dvou nebo více jednoduchých znaků MDT pomocí symbolů přiřazení či rozšíření nebo pomocí dvojtečky, dvojité dvojtečky či hranatých závorek

- Možnosti vytváření složených znaků spojovacími symboly

+ symbol přiřazení čte se „a“ nebo „plus“

Příklad: 622 +669 Hornictví a hutnictví

: symbol vztahu čte se „dvojtečka“
nebo „ku“

Příklad: 622.333:622.271 Těžba kamenného uhlí
povrchovým způsobem

/ symbol rozšíření čte se „až“ „lomeno“

Příklad: 626/627 Vodní stavitelství

:: symbol dvojité dvojtečky čte se „dvojitá
dvojtečka“

používá se k určení pevného, nezaměnitelného
pořadí (dnes již nepoužíván)

Symbol vztahu : „dvojtečka“
+ „plus“

může spojovat dva i více znaků

je možno je obrátit nebo určit prioritní znak

Příklad: 504.5:502.51 Znečištění hydrosféry

502.51:504.5

Spojování více než dvou znaků do vztahu

použijeme

symbol podřazení [] čte se „hranaté
závorky“

Příklad: [656:504]:34 Právní aspekty působení
dopravy na životní prostředí

Rozvinuté znaky MDT

Tvoří se připojením jednoho nebo více pomocných znaků k hlavnímu znaku MDT

pomocné znaky : **všeobecné**
zvláštní

Všeobecné pomocné znaky

Jsou použitelné v celé šíři MDT, pokud jejich spojení s hlavním znakem dává smysl

Všeobecné pomocné znaky jazyka

= čte se „rovná se“

Příklad: =162.3 čeština

Použití především při indexaci jazykových slovníků

Příklad: 81`374.8=111=112.2

anglicko – německý jazykový slovník

Všeobecné pomocné znaky formy

vyjádřeny nulou a dalšími číslicemi v
kulaté závorce

(0...) čte se "závorka nula"

Příklad:

(035) příručka

696.1(035) příručka instalátérských
prací

(075.3) středoškolské učebnice

51(075.3) středoškolská učebnice
matematiky

Všeobecné pomocné znaky místa

Vyjadřují se číslem, jehož první číslice je jiná než nula.

Číslo je uvedeno v kulaté závorce

(1/9) čte se „závorka„

Příklad: (437.3) Česko

(567) Irák

Všeobecné pomocné znaky času

“ “ čte se uvozovky

Příklad: “ 321“ Jaro

Všeobecné pomocné znaky etnické

(=...) čte se závorka rovná se

Příklad: (=414) Černoši obecně

(=622.82) Polynésané

Všeobecné pomocné znaky s pomlčkou a nulou

-0 čte se „pomlčka nula“

-03

pro vyjádření materiálu

Příklad:-035.35 Korek

-05

pro vyjádření osob

Příklad:-053.31 Novorozenci

Všeobecné pomocné znaky hlediska

.00 čte se“ tečka nula nula“

v nejnovějším vydání MDT se již

neuvádí

Zvláštní pomocné znaky

Uvádějí se přímo u některých jednotlivých oddílů, pro které platí jejich použití

Příklad: 62-5 Provoz strojů. Řízení a ovládání
66.014 Chemické složení

- Tabulky MDT
- Publikace nebo databáze , které obsahují seznamy třídníků a odpovídajících názvů v různém rozsahu a výběru

úplné vydání

zahrnuje všechny platné znaky MDT v

plném rozsahu

střední vydání

zkrácené vydání

speciální (oborové) vydání

- Současné vydávání MDT
- Konsorcium pro MDT (sídlo v Haagu)
 - převzalo agendu MDT od Mezinárodní federace pro dokumentaci
 - odpovědné za změny a doplňky
 - říjen 2009 Haag The International UDC Seminar 2009
 - využití MDT v automatizovaných systémech vyhledávání

ČR zastupuje v Konsorciu Národní knihovna ČR:

- výhradní licence na překlad a rozšiřování UDC Master Reference File (MRF) v ČR
- vydává MDT na CD-ROMU společně s firmou Albertina

nejnovější verze 5.vydání MDT na CD-ROMU

nové třídy v MDT: 004.Počítačová věda. Výpočetní technika
005 Management. Řízení
60 Biotechnologie

Vyhledávání MDT na CD ROMU

Globální index (vyhledává i v poznámkách):

Shift + F6

Zařazení do hierarchické struktury :

ikona otevřené knihy

- MDT na webu Národní knihovny ČR
5. vydání
stejný rozsah jako MDT na CD ROMU:

<http://aip.nkp.cz/mdt>

• KONSPEKT

vznik 70.léta

vytvořila skupina amerických knihoven pro svoji potřebu

hrubé rozčlenění informačních zdrojů

aplikace: - knihovní sbírky
- informační brány

cíl: koordinovaná tvorba fondů

předpoklad: - tématická struktura fondů

24 základních tématických skupin

dále se člení v druhé hierarchické úrovni na cca 5.000 kategorií

- Konkordance s DDC a MDT
 - jednotný standard pro popis a zpřístupnění informačních zdrojů
 - mezinárodně používaný propojovací jazyk
 - nástroj pro tvorbu předmětově orientovaných metainformačních systémů v mezinárodním prostředí
 - nástroj pro konverzi lokálních klasifikačních systémů do univerzálních klasifikačních systémů

–Konspekt na webu Národní knihovny

<http://konspekt.nkp.cz/konspekt.phtml>

Library of Congress Classification LCC (třídění Kongresové knihovny)

hlavní účel : organizace sbírek Kongresové knihovny
enumerativní systém –každé téma je explicitně
vyjádřeno

pomocné tabulky v menším rozsahu než u jiných
klasifikací

kombinace písmen latinské abecedy a arabských
číslic

např. : Antibiotika TP248.65.A57

využíváno pro třídění ve volném výběru NTK

LCC

skládá se:

- **main class** reprezentují vědní obory jednotlivých disciplín

např. Technology T

- **subclass** - označují se jedním nebo dvěma písmeny

např. Bridge engineering TG

- **division** History 15 -20
- **subdivision** 19.stol. 19

dějiny stavby mostů 19.stol. : TG19

LCC zohledňuje formu dokumentu,
čtenářské určení (učebnice , časopisy,
biografie, literatura pro mládež atd.)

např. TG9 Encyklopedie mostního
stavitelství

rok vydání dokumentu

zeměpisné určení

např. TG25.N53 New York. Brooklyn
Bridge

u každé třídy jiné

Signatura LCC

alfanumerické vyjádření jednotky v knihovním fondu

- **class number** – alfanumerický kód (tabulky)
- **item number** – odlišuje jednotky v rámci stejné třídy

(vytváří katalogizátor dle pravidel)

Cutterovo číslo (podle hlavního záhlaví)

- dle osobního jména
- dle názvu korporace
- dle názvu díla

rok vydání

číslo části

označení multiplikátu

- Cutterovo číslo :
přiřazení autorských znaků dokumentům pro
zohlednění abecedního stavění a řazení podle
systému Cuttera
použití tzv. Cutterovy tabulky
vytváří katalogizátor
může být v každé knihovně jiné (záleží na složení a
rozsahu fondu)

Např.: Kalina, Pavel: Dějiny středověké architektury. 2005

NA350 Architektura – historie – středověká architektura
- všeobecná díla

.K35 Cutterovo číslo

2005 rok vydání

Signatura: NA350.K35 2005

Fasetové třídění

S.R. Ranganathan (Madras, Indie)

Dvojtečková klasifikace (Colon Classification)

snaha - vyjádřit složená téma

(v dokumentech i uživatelských dotazech)

- redukovat enumerativní třídící tabulky

kombinační princip

1) tabulky základních tříd

2) tabulky , které reprezentovaly určitou charakteristiku předmětové oblasti

fasety

Předmětová hesla dle ČSN 010188

Zásada substantivní inverze

= na prvním místě podstatné jméno
následuje přídavné jméno

- vhodné pro lístkové katalogy
- současná generace OPACů umožňuje vyhledávat kterékoliv slovo z řetězce předmětového hesla
 - substantivní inverze ztrácí svůj význam

- Předmětové heslo se skládá z jednoho nebo více prvků

První prvek předmětového hesla - heslo

může být v nezbytných případech

upřesněno doplňkem hesla, který se uvádí
v kulatých závorkách

Příklady: Cukrovka (řepa)

Cukrovka (nemoc)

Praděd (likér)

Další prvky – podhesla

mohou být upřesněna doplňky

- Hesla a podhesla se navzájem oddělují pomlčkou s mezerou
- První slovo hesla se píše s velkými začátečními písmeny
- Podstatná jména se zpravidla uvádějí v 1. pádě množného čísla

Příklady :

Stromy ovocné – kmeny – ochrana

Automobily – konstrukce – učebnice

Odkazy

vylučovací odkazy :

odkazují od slov nebo slovních spojení , která nejsou v předmětových heslech využita na slova (slovní spojení) , kterých se využívá

před tvar na který se odkazuje se uvádí :

viz nebo zkratka v.

vylučovací odkazy odkazují většinou od jednoho synonyma na druhé

přidružovací odkazy:

navzájem spojují souběžně užitá slova nebo slovní spojení částečně synonymní nebo příbuzného obsahu

před tvar na který se odkazuje se se uvádí:

viz též

Klíčová slova

- slovo nebo slovní spojení vybrané z dokumentu a používané pro indexování

rozšířené pojetí

- nemusí být vybráno přímo z dokumentu

může být volně přiřazeno

vznik 50.léta – nové změny v technologii, nové dokumenty mající „víceoborový charakter“

1) automatická extrakce

(rejstříky KWIC a KWOC)

2) metoda intelektuálního výběru klíčových slov indexátory

Často jako pomocný prostředek, doplněk jiných selekčních jazyků (např. užití ve Státní technické knihovně)

TEZAURY

- jednojazyčné
- vícejazyčné

tezaurus - řízený a měnitelný slovník lexikálních jednotek, vyjadřující sémantické vztahy mezi lexikálními jednotkami a určený ke zpracování a vyhledávání informací

- **Deskriptor** - lexikální jednotka tezauru, která se používá ke zpracování a vyhledávání informací
- **Nedeskriptor** - lexikální jednotka tezauru, která se při zpracování a vyhledávání informací nahrazuje deskriptorem

- sice není dokumentu přiřazována, ale slouží jako uživatelský vstup do tezauru nebo abecedního rejstříku

uživatel je odkazován

slovem **viz** nebo zkratkou **ekv** k příslušnému deskriptoru

- výběr lexikálních jednotek tezauru závisí na volbě tématické oblasti, kterou informační systém pokrývá

- slovní spojení přídavné jméno - podstatné jméno se mohou uvádět v přirozeném nebo obráceném slovosledu

odkazy: vztah ekvivalence (synonymie)

vztah hierarchie (vztahy nadřazenosti

a

podřazenosti)

vztah asociace

Předmětová hesla Národní knihovny ČR (PHNK)

- přirozený pořádek slov
 - upustilo se od substantivní inverze
- tvořena podle pravidel tvorby předmětových hesel Kongresové knihovny

LCSH(= Library of Congress Subject Headings)

- nejrozšířenější a nejpropracovanější typ předmětového selekčního jazyka
- současné on-line katalogy jsou víceméně založeny na principech tohoto selekčního jazyka

PHNK

- původně prekoordinovaný selekční jazyk , v současné době modifikován prvky postkoordinace
- jednoduché řetězce
- hlavní nosná informace se musí vyjádřit ve vstupním prvku

Pořadí:

vstupní prvek tématický - tématické zpřesnění

- geografické zpřesnění – chronologické zpřesnění

Vztahy:

vztah ekvivalence – řeší vztahy mezi synonymy a kvazisynonymy, na jeden preferovaný termín jsou směřovány (odkazovány) jeho synonymické varianty

vztah hierarchie – stanovení nadřazených a podřazených pojmů

vztah asociace

Polytematický strukturovaný heslář (PSH)

- řízený slovník , určený pro věcný popis dokumentů v polytematických fondech a pro vyhledávání dokumentů
- obsahuje přes 16000 hesel v češtině a angličtině
- hesla jsou zařazena do 42 tématických řad (připravili specialisté pro jednotlivé obory)
- v řadě Obecnosti jsou zařazeny pojmy pro procesy a jevy, použitelné v různých odvětvích a oborech, které v kombinaci s konkrétními pojmy umožňují popis dokumentů

- každé heslo je podle obsahu a míry obecnosti zařazeno do šesti ,výjimečně sedmistupňové hierarchie
 - stromová struktura – od nejobecnějších pojmů až po pojmy užšího charakteru
 - každé heslo se vyskytuje pouze jedenkrát
 - PSH neobsahuje hesla geografická, chronologická, personální, názvy produktů, názvy organizací a profesních skupin
 - systém vylučovacích a přidružovacích odkazů

PSH

původní produkt Státní technické knihovny
(práce zahájena v roce 1993)

určen pro implementaci do automatických
knihovnických systémů

průběžně probíhá aktualizace hesláře

k dispozici dvojjazyčná verze 3.0

připravuje se verze 4.0

<http://pshmanager.techlib.cz/>

více informací :

lenka.satavova@techlib.cz

Přehled polí (UNIMARC a MARC 21) určených k zápisu jednotlivých selekčních prvků

osobní jméno použité jako předmět	600/UNI	600/MARC
jméno korporace použité jako předmět	601/UNI	610/MARC 611/MARC
jméno rodiny použité jako předmět	602/UNI	600/MARC 1.ind.3/MARC
jméno/název použité jako předmět	605/UNI	630/MARC
název použitý jako předmět	604/UNI	600/MARC 610/MARC 611/MARC
věcné téma použité jako předmět	606/UNI	650/UNI

geografické jméno použité jako předmět	607/UNI	651/MARC
forma, žánr (formální deskriptor)	608/UNI	655/MARC
volně tvořené termíny (klíčová slova)	610/UNI	653/MARC
předmětová kategorie (Konspekt)	615/UNI	072/MARC
MDT	675/UNI	080/MARC
kód časového období	661/UNI	045/MARC
LCC		050/MARC

Standardizace

- jednotnost popisu

předpoklad kooperace, sdílené katalogizace

Národní autority NK ČR

báze věcných autorit :

Věcné téma

Geografický termín

Skupina Konspektu

Formální deskriptor

MDT

Národní knihovna ČR –Katalogy a databáze – Další databáze NK ČR
Národní autority (AUT)

Nové trendy v popisu

elektronické dokumenty

- sémantický web

 - ontologie (vyjádření pojmů a vztahů mezi nimi)

- značkovací jazyky

 - (SQML)

- využití klasifikací pro prostředí webu

 - vyhledávání zdrojů na webu a jejich klasifikace

 - využití selekčních jazyků pro automatizované systémy vyhledávání

Literatura :

Kovář, B.: Věcné pořádání informací a selekční jazyky. Praha, ÚVTEI 1981. 144 s.

Balíková, M.: Změna koncepce věcného zpracování v Národní knihovně ČR. Praha, Národní knihovna 1998. 22 s.

Pinkas, O.: Zpracování informačních fondů. Praha, VŠE 2002. 206 s.

Macháčková, T.: Library of Congress Classification.

Děkuji za pozornost

Milada.Piskova@techlib.cz