

Téma 3 - Indexace

Indexace může znamenat:

- tvorbu indexu (indexového souboru)
- vyjadřování obsahu (věcný popis / zpracování) informací
⇒ **obsahová metadata**

ÚISK – OBA T03

1

1

Index může znamenat:

*latinsky: ukazovatel, ukazatel, udavatel, udavač, oznamovatel, nápis
digitus index = ukazováček (angl. index finger)*

- vyzvědač, zrádce, prozrazující něco
- seznam, obsah, katalog, rejstřík
- ukazatel poměru dvou nebo více hodnot
- rozlišující, identifikační znak; typ znaku (Peirce)
a mnoho konkrétních aplikací...

Indexace může znamenat:

- tvorbu indexu (indexového souboru)
- vyjadřování obsahu (věcný popis / zpracování) informací

ÚISK – OBA T03

2

2

Fasetový přístup k indexaci

Proces

- Tvorba indexu / indexového souboru

ČSN ISO 999

- Popis obsahu zdrojů / tvorba obsahových metadat ✓

ČSN ISO 5963

Zdroj

- Znak ✓
= slovní indexace

- Pojem ✓
= pojmová indexace

Granularita

- Makro-indexace ✓

- Mikro-indexace

ÚISK – OBA T03

3

3

Indexace jako popis obsahu zdrojů

ÚISK – OBA T03

4

4

Proces indexace podle ISO 5963

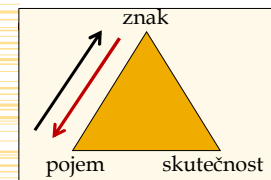
- 1) Analýza požadavku / dokumentu a určení jeho věcného **obsahu / předmětu**
- 2) Identifikace hlavních **pojmu** obsažených v předmětu
- 3) Vyjádření těchto pojmů **lexikálními jednotkami** selekčního jazyka

ÚISK – OBA T03

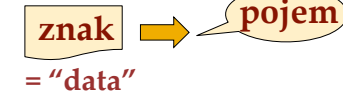
5

5

Proces indexace podle ISO 5963



- 1) Obsahová analýza



- 2) Vyjádření obsahu dokumentu nebo informačního dotazu prostřednictvím slov přirozeného jazyka nebo znaků umělého jazyka



ÚISK – OBA T03

6

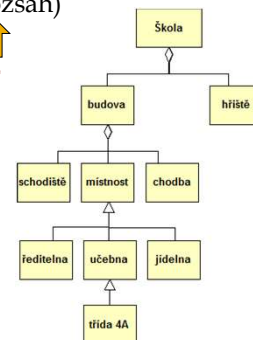
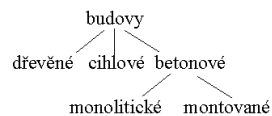
6

Co je to pojem?

Jednota intenze (obsah) a extenze (rozsah)

↑
↑
inverzní poměr

Zjednodušené chápání obsahu:
množina charakteristik
(vlastností, atributů)



ÚISK – OBA T03

7

7

Co je to lexikální jednotka? též **lexém**, *lexical unit*, *token*

obecně

- sekvence písmen, číslic a dalších specifických znaků používaná pro označení určitého **pojmu**
- jednotka jazyka, která je nositelem významu nebo funkce
- z hlediska významového tvoří jeden nedělitelný celek

v organizaci znalostí

- slovo nebo slovní spojení ve formě buď
 - výrazu přirozeného jazyka, nejčastěji substantiva nebo substantivního spojení
 - notačního znaku

ÚISK – OBA T03

8

8

Obsahová analýza textu

56

konkrétního významu textu, tak k určení kontextu⁴¹, potom vytvoření předpokladů pro její anadné využití je důležité při budování digitální infrastruktury pro digitální kodikologii a bibliologii a její transformaci z katalogizace a bibliografie na dějiny knižní kultury, resp. dějiny kultury vůbec.

Analytický popis ve významu, jak jsem jej nyní naznačil, se ovšem zcela vymyká významu, který mu přikládalo tradiční knihovnictví: není totiž použitým detailním rozpisem obsahu, nýbrž mnohem spíše vytvořením předpokladů pro postavení kontextuálních vztahů ve velkém bohatství a rozmanitosti, protože postihuje přinejmenším čtyři úrovně, totiž celek sbírky, resp. rukopisu, základní intelektuální jednotku, dílčí mikrotext a jednotlivý motiv⁴². Teoretický koncept takové vrstevnatosti je vcelku jasný a v obecném smyslu asi nenapadnutelný, nicméně velký deficit zůstává v jeho praktickém rozpracování jak v běžné katalogizaci a ediční práci, tak také v přesně vymezených monografických a analytických studiích. Vzhledem k tomu, že kontextuální vrstevnatost je v píni míře uplatnitelná pouze v síťovém digitálním prostředí, zatímco v tradičním tištěném prostředí a jeho konvenčních metodami je pravděpodobně zcela neuplatnitelná, je nutné její zprůtomňování vázat na rozvoj metodologie a technologie hypertextu⁴³, jež je nejvíce vhodnější vyjadřovací formou tohoto informačního a komunikačního prostředí. Zdálná však být zřejmé, že katalogizace, má-li být chápána jako „praktická kodikologie“, se v rámci digitální kodikologie výrazně odlišuje od editologie, což je pro její generace kodikologů s jejich fóbií z filologie zajiště iritující. Přece však se o tom nemějí diskutovat na širokém knihovnickém fóru⁴⁴. Otevírá se tudíž přímá cesta od popisu obsahového záznamu k pínému textu, od metadat k datům.

Knižní šifra

text – originál → **text**

ÚISK – OBA T03

9

Obsahová analýza textu

56

konkrétního významu textu, tak k určení kontextu⁴¹, potom vytvoření předpokladů pro její anadné využití je důležité při budování digitální infrastruktury pro digitální kodikologii a bibliologii a její transformaci z katalogizace a bibliografie na dějiny knižní kultury, resp. dějiny kultury vůbec.

Analytický popis ve významu, jak jsem jej nyní naznačil, se ovšem zcela vymyká významu, který mu přikládalo tradiční knihovnictví: není totiž použitým detailním rozpisem obsahu, nýbrž mnohem spíše vytvořením předpokladů pro postavení kontextuálních vztahů ve velkém bohatství a rozmanitosti, protože postihuje přinejmenším čtyři úrovně, totiž celek sbírky, resp. rukopisu, základní intelektuální jednotku, dílčí mikrotext a jednotlivý motiv⁴². Teoretický koncept takové vrstevnatosti je vcelku jasný a v obecném smyslu asi nenapadnutelný, nicméně velký deficit zůstává v jeho praktickém rozpracování jak v běžné katalogizaci a ediční práci, tak také v přesně vymezených monografických a analytických studiích. Vzhledem k tomu, že kontextuální vrstevnatost je v píni míře uplatnitelná pouze v síťovém digitálním prostředí, zatímco v tradičním tištěném prostředí a jeho konvenčních metodami je pravděpodobně zcela neuplatnitelná, je nutné její zprůtomňování vázat na rozvoj metodologie a technologie hypertextu⁴³, jež je nejvíce vhodnější vyjadřovací formou tohoto informačního a komunikačního prostředí. Zdálná však být zřejmé, že katalogizace, má-li být chápána jako „praktická kodikologie“, se v rámci digitální kodikologie výrazně odlišuje od editologie, což je pro její generace kodikologů s jejich fóbií z filologie zajiště iritující. Přece však se o tom nemějí diskutovat na širokém knihovnickém fóru⁴⁴. Otevírá se tudíž přímá cesta od popisu obsahového záznamu k pínému textu, od metadat k datům.

pojem

text – originál → **text**

ÚISK – OBA T03

10

Indexace textu

znak

pojem

znak

interakce

vyjádření

znak

text - originál

ÚISK – OBA T03

11

(Obsahová) metadata jako komunikační infra- nebo hyper-struktura

1

infra = dole, vespod, pod
hyper = nad, přes, nadmíru

1)

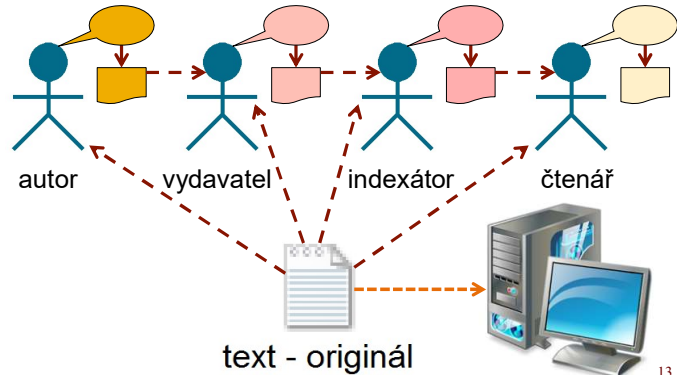
2)

3)

ÚISK – OBA T03

12

Faktory obsahové analýzy: subjekt, kontext, účel



13

Typy a vlastnosti „obsahové“ indexace

- pojmová
- slovní

- **úzká** ⇒ vyhledávání
- **široká/hrubá** ⇒ listování

ukazatele kvality:

- **úplnost** (*exhaustivity*)
- **přesnost** (*specificity*)
- **konzistence** ⇒ stejnorodost, předvídatelnost

ÚISK – OBA T03

14

14

Pojmová a slovní indexace

- **Pojmová – přidělování/přiřazování indexačních termínů**
assigned indexing

přiřazení obsahových metadat (bez ohledu na to, zda se vyskytují v textu) tak, aby reprezentovala informační zdroje podle určitých pravidel

Připomenutí:

SOZ i rejstříky organizují **pojmy**, nikoli slova!

- **Slovní – výběr/extrakce termínů**
derived, extracted indexing

postup vybírání indexačních termínů z indexovaného zdroje (termín ve zdroji ⇒ obsahová metadata)

ÚISK – OBA T03

15

15

Pragmatický aspekt indexace

„*warrant*“ = zdůvodněný, ospravedlněný, oprávněný

- **literary warrant**
- educational warrant
- user warrant
- organizational warrant
- business warrant

Příklad:

Kniha v TDKIV
Kniha v tezauru AGROVOC

„*bias*“ = zaujetí, zkreslení, předsudek, subjektivní ovlivnění

ÚISK – OBA T03

16

16

Indexace složených témat

Vyjádření vztahů mezi dílčími tématy:

- (1) Fasetová analýza
- (2) „Závislostní“ analýza
- (3) Předmětová hesla

ÚISK – OBA T03

17

17

Fasetová obsahová analýza = fasety jako kategorie

Obecné fasety/kategorie
aplikovatelné na velké množství typů obsahu

- věc (co)
- agent (kdo)
- proces
- událost
- materiál
- vlastnost
- místo (kde)
- čas (kdy)
- zpracování potravin
- vítěz závodu
- heterogenní síť
- porcelánová soška
- start raketoplánu
- doprava v Praze
- pondělní výuka
- karlovarský kaolin
- mistrovství Evropy

ÚISK – OBA T03

18

18

Struktura složených témat

(1) Souřadná

Sever a Jih, Čechy a Morava
Klasifikace a systémová analýza

(2) Podřadná

Jižní Čechy, moravské uzené
Etické faktory informačních interakcí člověka

ÚISK – OBA T03

19

19

„Závislostní“ analýza složených témat vytvořených podřadným skládáním

1. Určete významové jádro (*term focus*)
2. Určete rozlišující prvky (*term difference*)
3. Určete hledisko (*aspekt*), ev. žánr/formu či uživatelské určení – je-li aplikovatelné
4. Zformulujte termíny pro jednotlivé prvky

Příklad

Restaurování šelakových gramofonových desek

významové jádro	rozlišující prvek	hledisko
gramofonová deska	šelak (materiál)	restaurování

ÚISK – OBA T03

20

20

“Závislostní” analýza je syntagmatická

1. Synchronní a diachronní **výzkum** češtiny z hlediska strukturálního, kognitivního a funkčního
2. Financování základního **výzkumu** v AV ČR
3. Péče o talentovanou populaci v období docházky do mateřské a základní školy: průběžná **výzkumná** zpráva
4. Informační systém **výzkumu**, vývoje a inovací

významové jádro	rozlišující prvek	hledisko

ÚISK – OBA T03

21

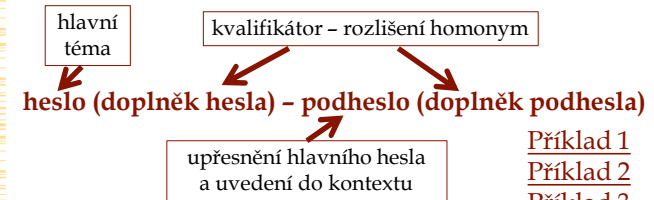
21

Předmětová hesla

ČSN 01 0188. *Tvorba předmětových hesel*. 1983

pravidla:

- nejužšího pojmu
- slovních spojení
- jazykové úpravy



ÚISK – OBA T03

22

22

Typy podhesel a pořadí jejich zápisu

1.	věcné upřesnění / zúžení významu hesla	<i>Akvizice ve veřejných knihovnách</i>	Knihovny veřejné - akvizice
2.	pojem ve vztahu k heslu (např. další téma)	<i>Veřejné knihovny a krajské spolky</i>	Knihovny veřejné - spolky krajské
3.	vztah hesla a podhesla typu 2	<i>Spolupráce veřejných knihoven a krajských spolků</i>	Knihovny veřejné - spolky krajské - spolupráce
4.	zeměpisné určení	<i>Spolupráce českých veřejných knihoven a krajských spolků</i>	Knihovny veřejné - spolky krajské - spolupráce - Česko
5.	časové určení	<i>Spolupráce českých veřejných knihoven a krajských spolků ve 21. století</i>	Knihovny veřejné - spolky krajské - spolupráce - Česko - 21. století
6.	forma/žánr	<i>Výzkum spolupráce českých veřejných knihoven a krajských spolků ve 21. století</i>	Knihovny veřejné - spolky krajské - spolupráce - Česko - 21. století - studie

ÚISK – OBA T03

23

23

Složené téma Problém = konstrukce třídníku/dotazu (tj. syntéza)

Možnosti vyjádření vztahů mezi dílčími tématy:
přirozený jazyk
umělý jazyk

1) **prekoordinace - enumerace složených témat**

2) **postkoordinace - booleovské operátory -**
OR, AND, (N), (W, PRE, -, “ ”)

ÚISK – OBA T03

24

24

Koordinace

souřadné spojení (syntagmatické)

využití **více lexikálních jednotek** k vyjádření **složených témat** (typ kolokace)

- **prekoordinace**
vztahy definovány už při tvorbě řízeného slovníku nebo při indexaci (na vstupu, apriori, „paradigmaticky“)
zvyšuje **přesnost**
- **postkoordinace**
vztahy definovány až při formulaci požadavku/dotazu (na výstupu, a posteriori, syntagmaticky)
zvyšuje **úplnost**

ÚISK – OBA T03 25

25

prekoordinace	
popis dokumentu	popis dotazu
<i>hodnocení profesorů studenty</i>	<i>hodnocení profesorů studenty</i>
postkoordinace	
popis dokumentu	popis dotazu
<i>hodnocení</i> <i>profesoři</i> <i>studenti</i>	<i>hodnocení</i> AND <i>profesoři</i> AND <i>studenti</i>

ÚISK – OBA T03 26

26

	prekoordinace	postkoordinace
klady	<ul style="list-style-type: none"> ■ přesné vyjádření tématu dokumentu (i velmi specifické vztahy) 	<ul style="list-style-type: none"> ■ možnost libovolného skládání pojmů
záporny	<ul style="list-style-type: none"> ■ obtížná formulace dotazu běžným uživatelem ■ nevyhovuje interdisciplinárním dotazům 	<ul style="list-style-type: none"> ■ omezení typů vztahů, jež lze použít

ÚISK – OBA T03 27

27

Prekoordinace vs. postkordinace

Vliv a účinky volného času na jednotlivce v České republice a na Slovensku v 21. století

<p>MDT</p> <p>379.8-042.3-059.1"20"(437.3+437.6)</p> <p>1 notační znak = 1 téma (prekoordinace)</p>	<p>Tezaurus EUROVOC</p> <p>volný čas Česká republika Slovensko sociální analýza 21. století</p> <p>1 téma = kombinace preferovaných termínů (postkoordinace)</p>
---	--

ÚISK – OBA T03 28

28

Prekoordinace v MDT

Vliv a účinky volného času na jednotlivce
v České republice a na Slovensku v 21. století
379.8-042.3-059.1"20"(437.3+437.6)

379.8	Volný čas. Dovolená, prázdniny	<i>hlavní znak</i>
-042.3	Vliv	<i>všeobecný pomocný znak vztahů, procesů a operací (postupů)</i>
-059.1	Osoby jednotlivě. Jednotlivé osoby	<i>všeobecný pomocný znak osob a osobních charakteristik</i>
"20"	21. století	<i>všeobecný pomocný znak času</i>
(437.3)	Česká republika	<i>všeobecný pomocný znak místa</i>
+	přiřazení	<i>všeobecný pomocný symbol</i>
(437.6)	Slovensko. Slovenská republika	<i>všeobecný pomocný znak místa</i>

29

Postkoordinace v JSTOR

□ Knowledge Organization and Text Organization

Suzanne M. Mannes, Walter Kintsch
Cognition and Instruction, Vol. 4, No. 2 (1987), pp. 91-115
JOURNAL ARTICLE

Topics: [Full text databases](#), [Advance organizers](#), [Inference](#), [Problem solving](#), [Paraphrase](#), [Bacteria](#), [Learning](#), [Memory retrieval](#), [Statistical significance](#)

□ Effects of Knowledge Organization on Task Performance

Bat-Sheva Eylon, F. Reif
Cognition and Instruction, Vol. 1, No. 1 (Winter, 1984), pp. 5-44
JOURNAL ARTICLE

Topics: [Physics](#), [History](#), [Problem solving](#), [Experiment design](#), [Performance tests](#), [Instructional materials](#), [Knowledge acquisition](#), [Educational psychology](#), [Cognitive models](#), [Memory](#)

30

30

Indexace jako tvorba indexu

- Indexový soubor
 - a) jako nástroj řešení lingvistických problémů obsahové analýzy a indexace
 - b) jako nástroj přístupu ke zdrojům

ÚISK – OBA T03

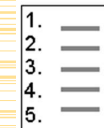
31

31

Metafory uspořádání (řazení) věcí v prostoru: seznam, kartotéka, index

přímý přístup

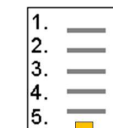
nepřímý přístup



seznam



kartotéka



indexový soubor

primární soubor

ÚISK – OBA T03

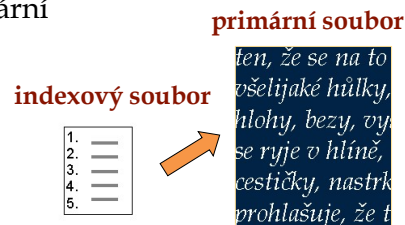
32

32

Indexy (rejstříky, registry)

dvoufázový přístup:

nejprve prohledáme index, pak primární soubor



ÚISK – OBA T03

33

33

Index

indexový soubor, invertovaný soubor, konkordanční soubor, rejstřík

- Pomocný soubor **řazený podle jiného hlediska** než primární (základní) soubor, obsahující odkazy na údaje v primárním souboru.

- obecná struktura: **klíč – adresa**

klíč:

hodnota vztahující se k primárnímu souboru (slovo, fráze, atribut) ⇒ **obsahová metadata**

adresa:

odkaz (ukazatel, pointer) na místo uložení v primárním souboru, jehož se hodnota týká ⇒ **lokační metadata**

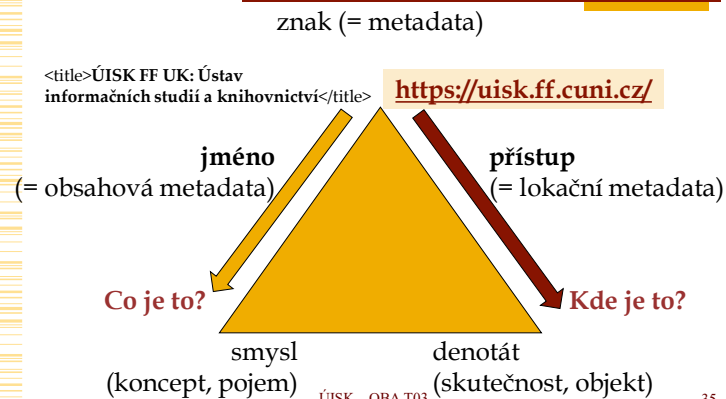
ÚISK – OBA T03

34

34

Princip indexu / rejstříku

Příklad: Web ÚISK FF UK (denotát)

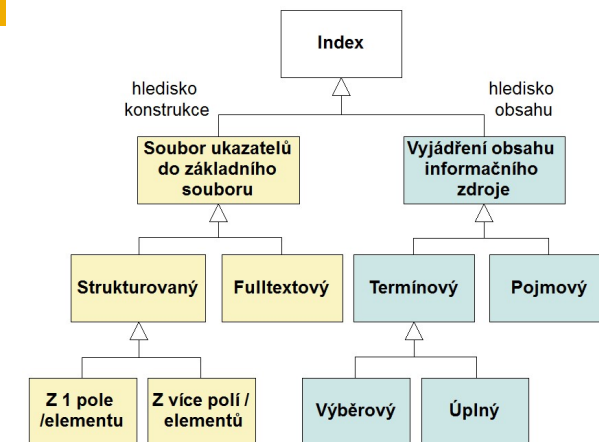


ÚISK – OBA T03

35

35

Typy indexů



36

36

Index – pro a proti

klad:

- urychluje přístup ke zdrojům (zkracuje dobu vyhledávání)
- umožňuje vyhledávat podle více hledisek

zápor:

- zabírá místo
- údržba indexu zpomaluje práci při aktualizaci primárního souboru

ÚISK – OBA T03

37

37

Obsahy publikací jako výběrové termínové indexové soubory

OBSAH	strukturní metadata	table of contents	lokační metadata
Úvod			7
1 Organizace			15
1.1 Historie			17
1.2 Problémy a paradigmat organizace znalosti			21
1.2.1 Konstantní problémy organizace znalosti			23
1.2.2 Nová paradigmat pro 21. století			26
1.3 Organizace znalostí jako obor			38
1.4 Pojmová základna a pojmová mapa organizace znalostí			46
1.5 Přístup k organizaci znalosti v tuzemském kontextu – poznámky (nejen) k české terminologii			51
1.6 Shrnutí			55
2 Systémový přístup k organizaci znalosti: cíl, prvky, procesy, kontext			57
2.1 Cíl organizace znalosti – proč se organizuje?			58
2.2 Kontext organizace znalosti – kdo, kdy, kde organizuje?			64
2.3 Prvky organizace znalosti – co se organizuje?			69
2.3.1 Jednotky organizace znalosti: dokument – informační zdroj – kolekce			70
2.3.2 Znalosti? Informace? Data?			77
2.3.3 Filozofický přístup k vymezení jednotky organizace znalosti			86
2.3.4 Vymezení jednotky organizace znalosti			92
2.4 Procesy organizace znalosti – co se děje při organizaci			94

Funkce

- Model (obsahové) struktury
- Nástroj přístupu ke strukturním prvkům

ÚISK – OBA T03

38

38

Rejstřík

Obchodním právu §§ 9 a 10, jakož i ve spise dra V. Kliea O zápisech do r-u obch. (Obch. knihovna »Merkur« č. 2). *Hmn.*

R. v knihách jest abecedně spořádaný seznam čili ukazatel jmen (**r.** jmenný), věcí (**r.** věcný) nebo slov (**r.** slovní), o nichž v knize buď se pojednává anebo jichž v knize bylo upotřebeno.

R. v knihtiskařství viz Punktura 2).

R., register, formát papíru, viz Papír,

Ottův slovník naučný
ilustrovaná encyklopaedie obecných vědomostí
Přejít na svazek knih
Část

Ottův slovník naučný
ilustrovaná encyklopaedie obecných vědomostí
21. R (R)-Rozkoš
Nakladatelské údaje
V Praze: J. Otto, 1904

ÚISK – OBA T03

39

39

Rejstříky publikací jako pojmové indexové soubory

Rejstřík

back-of-the-book index

obsahová metadata	lokační metadata	Funkce
A		
agregace 79		
akce 151		
změny ve verzích UML 150		
aktér 104		
aktivita 110		
aktivní třída 90		
aliasing 84		
Analysis Patterns 148		
analýza požadavků 46		
artefakt 101		
změny mezi verzemi UML 149		
asercce 64		
asociace 53		
časový signál 121		
část 132		
D		
datový pulec 73		
diagram		
členění typů 32		
začátky 35		
kompletní seznam 34		
spolupráce 137		
tříd 32		
typy 31		
úhel pohledu 28		

Funkce

- Model jednotek obsahu
- Nástroj přístupu k jednotkám obsahu

ÚISK – OBA T03

40

40

Struktura rejstříku publikace

generalizace 155
 viz též abstrakce; specializace
 - v UML 238-239
 generická hierarchie 1
 Genová ontologie 185
 genus-species 181
 globální číslo obchodní položky viz GTIN
 globální polohovací systém viz GPS
 gnozo... orie poznání
 Google..., 131, 136
 GLN (Global Location Number) 145
 GPS 141
 graf 28, 63, 167
 viz též teorie grafů
 - hrana viz hrana grafu
 - kořen viz kořen grafu

Základní struktura:
 obsahová a lokační metadata
 (klíč - adresa)

Vyjadřované vztahy:

- ekvivalence
- asociace
- hierarchie

lokační metadata (okolo 181)

obsahová metadata (okolo 141)

ÚISK - OBA T03 41

41

Předmětový rejstřík DDT Relative index

abecedně uspořádaný soupis témat vyjádřených
v přirozeném jazyce + odkazy na třídy

zahrnuje pojmové vztahy:

- vztah téma (*subject*) - obor (*discipline*)
- vztah asociace (*see also*)

ÚISK - OBA T03 42

42

Museum guidebooks	T1-074
Museum labels	069.53
Museum libraries	027.68
Museum...	069.1
Museum... methods	069.1
Museum science	069
Museum services	069.1
Museums	T1-075 5
	069
	T1-074
architecture	727.6
area planning	711.57
law	344.093
public administration	350.853
central governments	351.853
local governments	352.945 3

obsahová metadata (okolo 069)

lokační metadata (okolo 069)

subject (téma, předmět) → **discipline** (hledisko, obor)

ÚISK - OBA T03 43

43

Automation (continued)	303.483 4	Automobiles (continued)	388.342
social effects	629.809 2	transportation services	388.413 21
Automation engineers		urban	511.33
occupational group	T7-629 8	Automorphic functions	515.9
Automation training		calculus	512.7
personnel management	658.312 43	number theory	516.1
Automatons	629.8	Automorphisms	512.55
engineering	629.8	geometry	512.55
Automobile bodies	629.26	topological algebras	512.55
Automobile cars (Railroad)	385.34	Automotive industry	338.476 292
engineering	625.24	Automotive vehicles	388.34
see also Rolling stock		engineering	629.2
Automobile driving	629.283	international law	341.756 84
international law	341.756 86	law	343.094 4
law	343.094 6	military engineering	623.747
international law	629.209 2	production economics	338.476 292
Automobile engineers		public administration	350.878 34
occupational group	T7-629 2	central governments	351.878 34
Automobile industry		local governments	352.918 34
law	343.078 629 2	repair	629.287
Automobile insurance	368.092	safety engineering	629.204 2
inland marine	368.232	safety services	363.125
law	346.086 092	sports	796.7
liability	368.572	transportation services	388.34
see also Insurance		Autonomic nervous system	
Automobile noise	363.741	human anatomy	611.83
social welfare	363.741	human diseases	616.88
see also Noise		medicine	612.89
Automobile parking	388.474	human physiology	
see also Parking facilities		see also Nervous system	
Automobile racers	796.720 92	Autonomous authorities	
sports group	T7-796 7	public administration	350.009 1
Automobile racing	796.72	central governments	351.009 1
electronic games	794.867 2	local governments	352.009 2
Automobile rallies	796.73	Autonomy of states	320.15
Automobile registration	350.878 34	Autopilots	
international law	341.756 84	aircraft	629.135 2
law	343.094 4	Autopsy	
public administration	350.878 34	forensic medicine	614.1
Automobile safety	363.125	medicine	616.075 9
law	343.094 4	Autumn	
Automobile theft	364.162		

asociativní vztahy

ÚISK - OBA T03 44

44

Pojmové vyhledávání v bibli (konkordance – slovo v kontextu)

Strom — Střecha.

747

1 M. 1, 29. všel stromovi, na němž i ovoce s-u l Zi. 7, 3. neškodte zemi, ani moři, ani s-ům

Strana 1 z výsledků vyhledávání fráze "strom" (celkem 181 veršů - 10 stran)

Kliknutím na číslo verše se text zobrazí v kontextu kapitoly

lokační
metadata

Genesis 1, 11 Bůh také řekl: „zazeleněj se země zelení: bylinami, které se rozmnožují semeny, a ovocným stromovím rozmanitého druhu, které na zemi ponese plody se semeny!“ A stalo se tak.

Genesis 1, 12 Země vydala zeleň: rozmanité druhy bylin, které se rozmnožují semeny, a rozmanité druhy stromoví, které nese plody se semeny. Bůh viděl, že to je dobré.

Genesis 1, 29 Bůh také řekl: „vám na celé zemi každou bylinu nesoucí semena i každý strom, na němž rostou plody se semeny, dávám vám na pokrm.“

Genesis 2, 16 A Hospodin řekl: „Z každého stromu zahrady smíš jíst.

Genesis 2, 17 Ze stromu poznání dobrého a zlého však nejz. V den, kdy bys z něho pojedel, propadneš smrti.“

obsahová
metadata

2 Kř. 18, 17. položili se u s-y rybníka; 30, 2. Iz. 7, 3. konec s-y rybníka hořejšího
— 17, 10. obrazů pod každým s-em zeleným
Job 14, 7. o s-u j, naděje, že se zase zotaví
10-10. vyvrátil jako s- nadějí mou

Strana.

45

45

Rejstřík jako kolektivní dílo

A tak již nezbyvá než i na tomto místě vyslovit srdečné díky našim 250 spolupracovníkům při přípravných excerpčních pracích, a zvláště našim spolupracovníkům nejbližším, bez jejichž obětavé pomoci by Konkordance ani nemohla vycházet, totiž sestře Martě Hudcové, roz. Hanušové a bratru Jaroslavu Sekerkovi z Nasavrk u Chrudimě a sestře Martě Kocnové, roz. Hrejsově z Prahy. Jejich práce je zhodnocena v zmíněném již článku „Pět let Biblické konkordance“ v Theologické příloze Křesťanské revue, 1960.

Pokud pak jde o hlavní autory a vydavatele, budiž připomenuto, že M. Bič odpovídá za český text celého materiálu starozákonního a za ekvivalenty hebrejské a aramejské, J. B. Souček za český text novozákonní a řecké ekvivalenty starozákonní i novozákonní.

BIČ, Miloš, SOUČEK, Josef Bohumil, ed. *Biblická konkordance*. 1. vyd. Praha: Ústřední církevní nakladatelství, 1961. Díl 1 A-K, s. vii.

ÚISK – OBA T03

46

46

Systémy organizace znalostí a rejstříky – systémový přístup

Účel:

1) Pojmová indexace zdrojů
v určené doméně

2) Přístup ke zdrojům

Funkce:

Pomůcka při organizaci znalostí

Struktura:

Prvky: metadata

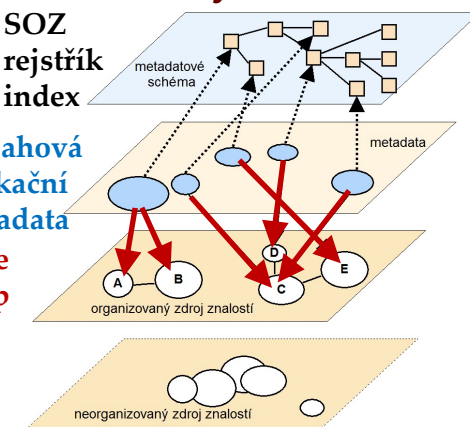
Vztahy: paradigmatické

ÚISK – OBA T03

47

47

Systémy organizace znalostí a rejstříky jako modely metadata

SOZ
rejstřík
indexObsahová
a lokační
metadataIndexace
a přístup

ÚISK – OBA T03

48

48

Systémy organizace znalostí a rejstříky jako řízené slovníky

Slovníky uspořádané tak, že explicitně zachycují
paradigmatické vztahy mezi pojmy.

Nástroj vyjádření vztahů: **notace**

33 Ekonomika. Ekonomické vědy

330 Ekonomika obecně

330.1 Ekonomická instrumentální hudba

generalizace 155

viz též abstrakce; specializace

- v UML 238-239

generická hierarchie 181-182

Genová ontologie 185

genus-species 181

globální číslo obchodní položky viz GTIN

globální polohovací systém viz GPS

alt BT zpěvní hlasy

baryton BT zpěvní hlasy

bas BT zpěvní hlasy

dechový nástroj **skos:broader** Nástroj

trunný nástroj **skos:broader** Nástroj

létna **skos:broader** Dechový nástroj

ousle **skos:broader** Strunný nástroj

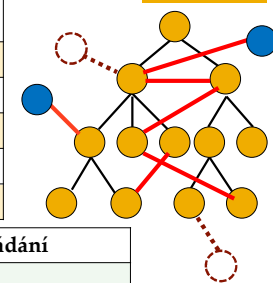
ousle **skos:related** Houslový koncert

ytara **skos:narrower** Akustická kytara

49

Systémy organizace znalostí a rejstříky – struktura

Název základní jednotky (=pojem)	
klasifikační schéma, ontologie	třída
tezaurus	termín
Wikidata	položka
rejstřík	záhlaví
index	klíč



Základní paradigmatické vztahy		uspořádání
klasifikační schéma, ontologie	hierarchie	hierarchické
tezaurus, rejstřík	ekvivalence	abecední (lineární)
Wikidata	asociace	síťové

UISK – OBA T03

50

50

Struktura tezaurového odstavce (record)

PREFERRED TERM	změna klimatu
DEFINITION	As alterações climáticas são mudanças do clima que ocorrem a nível global. Segundo o "Painel Intergovernamental sobre Mudanças Climáticas" (IPCC, 2001) as alterações climáticas resultam de uma variação estatisticamente significativa (no mínimo mais de 30 anos) no estado médio do clima ou na sua variabilidade, sem identificação absoluta das causas dessa mudança. (pt)
BROADER CONCEPT	přírodní jevy (cs) životní prostředí (cs)
NARROWER CONCEPTS	anthropogenic climate change (en) globální oteplení (cs)
RELATED CONCEPTS	adaptace na změnu klimatu (cs) climate system (en) zmírění dopadů změny klimatu (cs)
ENTRY TERMS	změny klimatu (cs)
INFLUENCES	klima (cs) narušování ozonové vrstvy (cs)
IS INFLUENCED BY OR DEPENDS ON	obohacení kyslíčkem uhlíčitým (cs) skleníkové plyny (cs)

51

51

Struktura rejstříkového hesla (index entry)

1. záhlaví (heading)
- 2a) **lokátor/y** (locator/s) záhlaví
- 2b) vylučovací odkaz
("see" cross-reference)
3. přidružovací odkaz
("see also" cross-reference)
4. podzáhlaví (subheading)
- 5a) **lokátor/y** (locator/s) podzáhlaví
- 5b) vylučovací odkaz
6. přidružovací odkaz

UISK – OBA T03

generalizace 155

viz též abstrakce; specializace

- v UML 238-239

generická hierarchie 181-182

Genová ontologie 185

genus-species 181

globální číslo obchodní položky

globální polohovací systém viz

gnozeologie viz teorie poznání

Google 11, 36, 103, 131, 136

GLN (Global Location Number)

GPS 141

graf 28, 63, 167

viz též teorie grafů

- hrana viz hrana grafu

52

Rekapitulace

Funkční požadavky na bibliografická metadata dle IFLA LRM

- Najít (*find*)
- Určit (*identify*)
- **Vybrat (*select*)**
- Získat (*obtain*)
- **Prozkoumat (*explore*)**

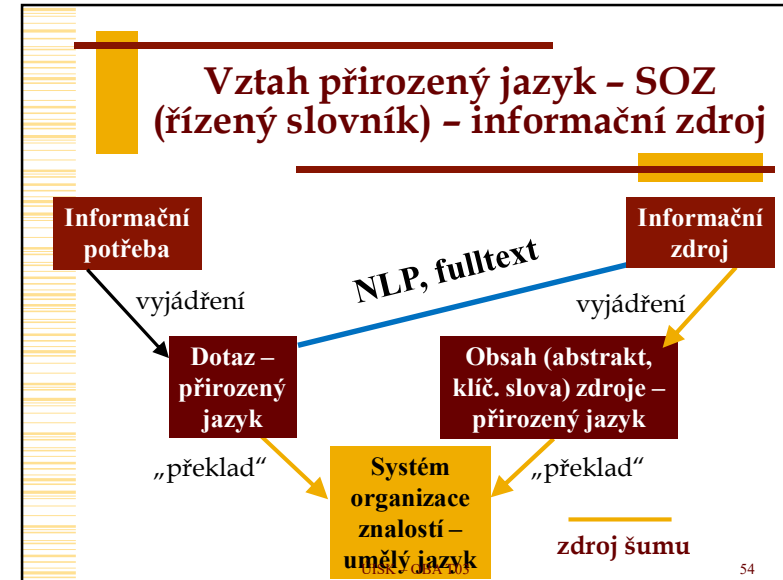
} informační zdroj

1) Intelektuálně vytvořená ✓
2) Automaticky vytvořená ✓
3) Bez metadat? ✓

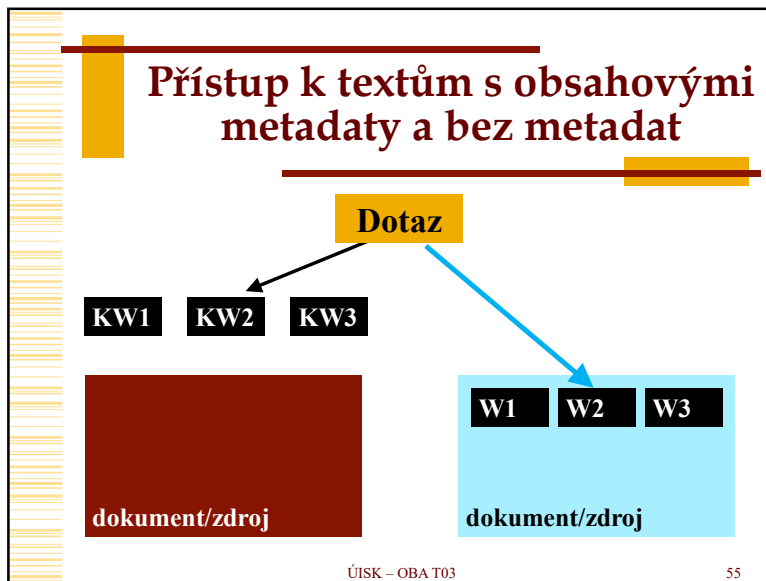
Obsahová metadata

ÚISK – OBA T03 53

53



54



55

Přístup s obsahovými metadaty – pro a proti

- + vyřešená synonymie a homonymie
- + nezávislost na slovníku autora dokumentu (v metadatech je možné popsat obsah i slovy, jež nejsou v dokumentu)
- + k dispozici je model (abstrakce) obsahu, zbavený detailů
- + umožňuje pojmové vyhledávání
- + možnost interpretace, hodnocení
- vyžaduje intelektuálně předzpracovat dokumenty i dotazy
- v bodech zpracování dochází k šumu

ÚISK – OBA T03 56

56

Přístup bez obsahových metadat - pro a proti

- + dokumenty ani dotazy není třeba předem intelektuálně zpracovat
- + ubývá bodů zpracování (tj. zdrojů šumu)
- závislost na slovníku autora dokumentu
- problémy s přirozeným jazykem (synonymie, homonymie, vágnost)

ÚISK – OBA T03

57

57

Jak dostat to nejlepší z obou světů?

Přístup bez metadat:

- Místo porovnávání stručného textu dotazu s rozsáhlým textem dokumentu – statistické porovnávání (celého) textu relevantního zdroje s texty v prohledávaném souboru (např. při detekci spamu)

Přístup s metadaty:

- Místo intelektuálního vytváření metadat – automatické generování metadat, automatická indexace

ÚISK – OBA T03

58

58

Vyhledávání na Internetu

lokační metadata

obsahová metadata

slovo (klíč)

http://www... pes

indexový soubor

primární soubor

ÚISK – OBA T03

59

59

Tzv. fulltextový (volný, úplný) index

- užití: narativní dokumenty, vyhledávání v textech
- *information retrieval*: celý dokument je rozložen na jednotlivá slova (lexikální jednotky)
- vyhledávaná jednotka: text (dokument), příp. jeho část
- složení:
 1. slovo (lexikální jednotka)
 2. identifikátor dokumentu v kolekci
 3. identifikátor elementu dokumentu (je-li dokument členěn na elementy)
 4. pořadí slova v rámci dokumentu (příp. elementu)

ÚISK – OBA T03

60

60

Tzv. strukturovaný index

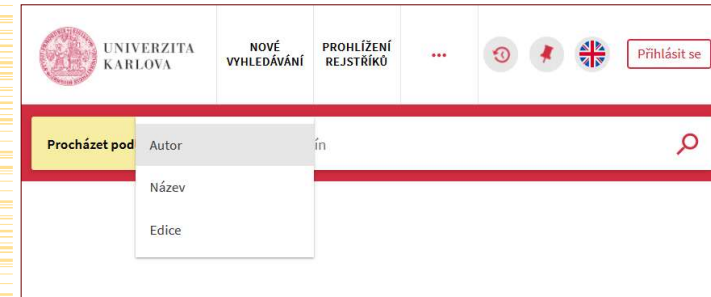
- užití: transakční dokumenty, vyhledávání v metadatech
- *database searching*
index s celými hodnotami elementů / položek
- vyhledávaná jednotka: záznam
- složení:
 1. hodnota pole (elementu)
 2. pořadové číslo záznamu v souboru
 3. pořadové číslo pole (elementu) v záznamu

ÚISK – OBA T03

61

61

Strukturovaný index vytvořený z hodnot jednoho pole (elementu)



ÚISK – OBA T03

62

62

Strukturovaný index vytvořený z hodnot jednoho pole (elementu)

Primární soubor

Indexový soubor

Hodnoty (klíče)	Číslo záznamů	Ukazatele	Hodnoty (klíče)
AUTOR	NÁZEV		NÁZEV
Čapek	Loupežník	118	Apollinaire
Apollinaire	Pražský chodec	117	Idiot
Hartwigová	Apollinaire	118	Inzerát na dům, ve kterém už nechci bydlet
Čapek	R.U.R.	119	Kaligramy
Čapek	Krakatit	120	Krakatit
Dostojevskij	Idiot	121	Loupežník
Čapek	Věc Makropulos	122	Ostře sledované vlaky
Hrabal	Pábitelé	123	Nikola Šuhaj loupežník
Čapek	Život a dílo skladatele Foltýna	124	Pábitelé

63

Strukturovaný index vytvořený z hodnot více polí (elementů)

- a) hodnoty různých polí (elementů) jsou umístěny do stejného pole indexového souboru
- b) složený index
původní položky zůstanou zachovány, třídí se kaskádovitě

ÚISK – OBA T03

64

64

Strukturovaný index z hodnot více polí – varianta a)

Základní vyhledávání | Rozšířené vyhledávání | Z více bází | Vyhledávání CCL

NKC - Online katalog Národní knihovny ČR

[Upozornění na dočasné omezení půjčování](#)
[Tematická mapa fondů](#) [Informace o bází](#)

Základní vyhledávání

Vyberte údaj pro vyhledávání:

Zadejte slovo nebo slovní spojení:

Druh dokumentu: pouze dokumenty v češtině

Blížkost slov? Ne Ano

Pro další informace použijte tabulku [kódů jazyků](#), tabulku [kódů zemí](#) a tabulku [kódů druhů dokumentu](#).

ÚISK – OBA T03 65

65

Strukturovaný index z hodnot více polí – varianta b) složený index

Hodnoty (klíče)		Číslo záznamů (ukazatele)
AUTOR	NÁZEV	
Apollinaire	Kaligramy	127
	Pražský chodec	117
Čapek	Krakatit	120
	Loupežník	116
	R.U.R.	119
	Věc Makropulos	122
	Život a dílo skladatele Foltýna	124
Dostojevskij	Idiot	121
Hartwigová	Apollinaire	118
Hrabal	Inzerát na dům, ve kterém už nechci bydlet	126
	Ostře sledované vlaky	125

66

66

Využití indexového souboru pro kontextové operátory

operátor proximity

podmínka vzdálenosti se vyhodnocuje přes absolutní hodnocení rozdílu pořadí slov

operátor pořadí

podmínka pořadí je vyhodnocována na základě větší velikosti čísla udávajícího pořadí slov

ÚISK – OBA T03

67

67

Využití indexového souboru pro kontextové operátory

Text

Využití pro kontextové operátory

Index

2	indexového
5	kontextové
6	operátory
4	pro
3	souboru
1	využití

Dotaz

indexového (W) souboru
“indexového souboru”

ÚISK – OBA T03

68

68

Rotované rejstříky

KWIC a KWOC

KEYWORD-IN-CONTEXT BIBLIOGRAPHICAL INDEX

COLUMB EXCHANGE ENERGY FROM SHELL-MODEL NAV	1719
EXCITATION OF PROTONS IN HELIUM 11 B	0011
OF ATOMIC AND MOLECULAR EXCITATION BY A TRAPPED-ELECTRON IN E	0150
THERMAL EXCITATIONS IN LIQUID HE3.	1465
ENERGIES OF GROUND AND EXCITED NUCLEAR CONFIGURATIONS IN TH	0452
EXCITED STATES OF V51 AND CR53.	1891
4-PLUS EXCITED STATE IN OSMIUM-188.	1717
INTERNAL PHOTOEFFECT AND EXCITON DIFFUSION IN CADMIUM AND ZIN	0123
OF THE CONTRIBUTION OF EXCITONS TO THE COMPLEX DIELECTRIC	1555
THERMAL EXPANSION OF SOME CRYSTALS WITH THE	0136
ENERGY LEVELS IN F10 FROM THE N14/ALPHA,ALPHA/N14 AND	0547
ON FROM AL27-PLUS-P AND F19-PLUS-P.	0239
TIC MEASUREMENTS OF THE FE-CR SPINELS.	1603
BARIUM FERRATE III.	0326
MAGNETOSTATIC MODES IN FERRIMAGNETIC SPHERES.	0059
NICKEL-IRON FERRITE.	0397
TRANSITION TO THE FERROELECTRIC STATE IN BARIUM TITANA	0413
SUPERCONDUCTIVITY AND FERROMAGNETISM IN ISOMORPHOUS COPPOU	0089
INTERPLANETARY MAGNETIC FIELD AND ITS CONTROL OF COSMIC-RAY	0589
MAGNETIC FIELD DEPENDENCE OF ULTRASONIC ATTEN	0080
RELATIVISTIC FIELD THEORY OF UNSTABLE PARTICLES.	0283
QUANTUM FIELD THEORIES WITH COMPOSITE PARTIC	0669

69

Řešení přístupu k víceslovným výrazům pomocí abecedního seřídění konkordance

a jiných oprávněných podle jejich čísel , b) abecední rejstřík vlastníků a jiných oprávněných , c) soupis parcel podle	
tisku s možností jejich snadného vytištění , stejně jako abecední rejstřík pro snazší nalezení požadovaných informací , nastavení apod . Přímé	
řazení na CD-ROM a podrobný abecední rejstřík studijních oborů . Nabídka vzdělávání na veletrhu Gaudeamus je urč	
nti k dispozici podrobný abecední rejstřík oborů a bezplatný poradenský servis při vyhledávání žádaného studia	
levostranného ally měla v úmyslu projít abecední rejstřík držitelů patentů a zjistit , jestli tam něco je pod	
kontextu a znovu si vzala k ruce abecední rejstřík britských patentů . Jestliže chtěl Bellmann na té věci vydělat	
jisti , kde najít potřebné informace , pomůže vám abecední rejstřík hesel na straně 112 . Na stranách 14 - 15	
komplikovanějšími naučnými knihami . Na konci knihy je zařazen abecední rejstřík , ve kterém se děti mohou naučit hledat předmět ,	
katalog s širší nabídkou možností studia na CD-ROM a podrobný abecední rejstřík studijních oborů . Nabídka vzdělávání na veletrhu Gaudeamus je	
řazení mavnosti , například letos poprvé mají studenti k dispozici podrobný abecední rejstřík oborů . Nabídka vzdělávání na veletrhu Gaudeamus je	
koupil státní cenné papíry . Sally měla v úmyslu projít abecední rejstřík držitel	
pravostranného blematikám tisku s možností jejich snadného vytištění , stejně jako abecední rejstřík pro sn	
řazení , komplikovanějšími naučnými knihami . Na konci knihy je zařazen abecední rejstřík , ve kter	
kontextu nejste jisti , kde najít potřebné informace , pomůže vám abecední rejstřík hesel	
řazení , něco ji napadlo a znovu si vzala k ruce abecední rejstřík britských patentů . Jestliže chtěl Bellmann na té věci	
vlastníků a jiných oprávněných podle jejich čísel , b) abecední rejstřík vlastníků a jiných oprávněných , c) soupis parcel	
ÚISK – OBA T03	70

70