



VIKMA05

Organizace znalostí

5. 4. 2019: Formální struktura dokumentu, reprezentace dokumentu (P7+K3)

FF MU, jaro 2019

Mgr. Josef Schwarz

126172@mail.muni.cz



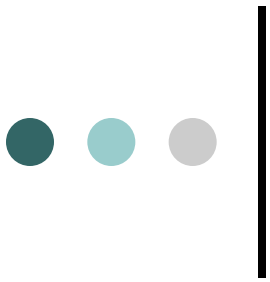
Osnova

- Praktická část
- Přehled (minulých a budoucích) témat
- Teoretická část
 - SGML
 - právní texty – struktura
- Zadání 7. dílčího kroku DA



Praktická část

DA krok č. 6



PŘEHLED TÉMAT



Přehled (minulých) témat

Témata

- Úvod
- Přístupy ke KO, doménová analýza
- Ontologie, epistemologie, teleologie (pojem, metafora)
- Kategorie, kategorizace, klasifikace (tradiční vs kognitivistický přístup: prototypy, bázové kategorie)
- Termíny a terminologie
- Notace a notační systém

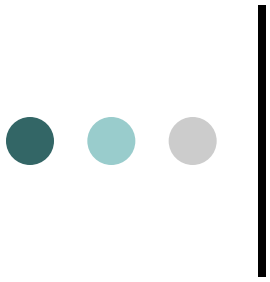
Úkoly

- Volba domény
- Zdroje
- Metafory
- Prototypy a bázové kategorie
- Reprezentace pojmu
- Návrh notačního systému



Přehled (budoucích) témat

- Formální struktura dokumentu, reprezentace dokumentu
- Procesy organizace znalostí
- Ontologie
- Univerzální klasifikační systémy
- Systémy založené na přirozeném jazyce
- Reprezentace a vizualizace znalostí
- Závěrečná hodina - test



STRUKTURA DOKUMENTU



Formální struktura dokumentu

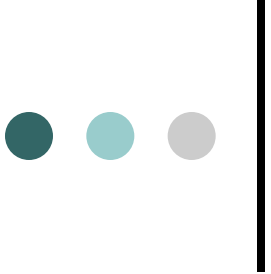
- Tradiční

- Obsah, rejstřík, marginálie, živé záhlaví aj.

- Elektronická

- SGML (ISO 8879:1986)

- Obecná struktura dokumentu a další atributy
- Odvozené jazyky: např. HTML, XML
- Pro základní informace o SGML a jeho struktuře viz web W3.org



SGML – příklad 1

PhDr. Rostislav Kaprál

Dělnická 356/5

280 02 Kolín II



SGML – příklad 1, varianta A

- <DOKUMENT>
 - <adresa stát="cs">
 - <osoba>
 - <titul>PhDr. </titul><křestní jméno>Rostislav</křestní jméno><příjmení>Kaprál</příjmení>
 - </osoba>
 - <ulice>Dělnická</ulice>
<číslo_popisné>356</číslo_popisné>/<číslo_orientační>5</číslo_orientační>
 - <PSČ>280 02</PSČ> <město>Kolín </město>
<městská část> II </městská část>
 - </adresa>
- </DOKUMENT>



SGML – příklad 1, varianta B

- <DOKUMENT>
 - <adresa stát="cs">
 - <subjekt typ="osoba">
 - <titul>PhDr. </titul><křestní jméno>Rostislav<křestní jméno><příjmení>Kaprál</příjmení>
 - </subjekt>
 - <veřejný_prostor typ="ulice">Dělnická</veřejný_prostor>
<číslo_popisné>356</číslo_popisné>/<číslo_orientační>5</číslo_orientační>
 - <PSČ>280 02</PSČ> <město>Kolín </město>
<městská část> II </městská část>
 - </adresa>
- </DOKUMENT>



SGML – příklad 2

- Nejvýznamnější a největší jeskyní Tišovského krasu je jeskyně Králova. Byla objevena v květnu roku 1972 náhodně v souvislosti s jezevčí norou. Dostala jméno po Aloisi Královi, objeviteli Demänovských jeskyní na Slovensku, který byl již v době objevu čestným občanem města Tišnova. Nejstarší objevené části Královy jeskyně mají délku 350 m a bohatou krápníkovou výzdobu. V letech 1974 a 1984 byly objeveny další části, pozoruhodný je Dobšiarův dóm s brčky dlouhými až 160 cm.

SGML – příklad 2

```
<dokument typ=„průvodce“ podtyp=„přírodní objekty“>
<objekt typ=„podzemní prostory“ podtyp=„charakteristika“>
<vnitřní prostora typ=„jeskyně“
podtyp=„historie“>Nejvýznamnější a největší jeskyní <lokalita
typ=„území“ podtyp=„kras“>Tišovského krasu </lokalita> je
jeskyně Králova. <událost typ=„objev“> Byla objevena v <čas
typ=„datum“ podtyp=„měsíc-rok“>květnu roku 1972 </čas>
<okolnosti>náhodně v souvislosti s jezevčí norou </okolnosti>
</událost>. Dostala jméno po <subjekt typ=„osoba“
podtyp=„speleolog“> Aloisi Královi, objeviteli Demänovských jeskyní na
Slovensku, který byl již v době objevu čestným občanem města
Tišnova</subjekt>. <rozloha typ=„délka“> Nejstarší objevené části
Královy jeskyně mají délku 350 m </rozloha> a <prvky
typ=„krápníky“> bohatou krápníkovou výzdobu </prvky>. <událost
typ=„objev“> <čas typ=„datum“ podtyp=„rok“> V letech 1974 a
1984 </čas> byly objeveny další části, pozoruhodný je <součást
typ=„dómy“> Dobšiarův dóm <prvky typ=„krápníky“>s brčky
<velikost typ=„délka“>dlouhými až 160 cm </velikost>. </prvky
</součást> </událost> </vnitřní prostora> </objekt>
</dokument>
```



Analýza právních textů

Kořenský – Cvrček – Novák: Juristická a lingvistická analýza právních textů. Praha: Academia, 1999

○ A. Typologie dokumentů

● 1. podle druhu

- **právní předpisy**
- **soudní rozhodnutí**
- **individuální rozhodnutí a normativní instrukce státní správy**
- **důvodové zprávy**
- **návrhy a projekty zákonů a dalších předpisů**
- **záznamy parlamentních debat**
- **výsledky připomínkových řízení**
- **komentáře**
- **učebnice, skripta a odborné monografie**
- **odborné články**
- **expertizy**
- **přehledy platných právních předpis**



Analýza právních textů

- **A. Typologie dokumentů**
 - **2. Podle formy publikace**
 - oficiální sbírky
 - neoficiální sbírky
 - knihy
 - odborné časopisy
 - jednotlivé dokumenty uložené v archivech institucí
 - **3. Podle vztahu k původnímu znění**
 - původní
 - upravené
 - obohacené
 - metadokumenty nebo adresy



Analýza právních textů

- **B. Struktura položek v základním slovníku (LEMMA)**
 - 1. Základní slovní tvar
 - 2. Typ
 - Speciální znaky
 - Číslice, římské číslice
 - Zkratky s tečkou (*Sb.*), zkratky bez tečky (*ČR*)
 - Pomocné znaky (§)
 - Slovenská slova
 - Cizí slova (ne chemické vzorce, ne názvy nemocí apod.)
 - 3. Slovní druh
 - 4. Rod
 - 5. Životnost
 - 6. Příslušnost ke kmenové sémantické skupině (*daně, zdanění, danit*)
 - 7. Typy právních významů
 - 8. Četnost
 - 9. Číslo základního slovního tvaru



Analýza právních textů

○ C. Sémantika morfologicko-syntagmatických typů

● Příklad – škoda

- *A – S:* *případná / věcná / hrozící*
- *S – S:* *velkého rozsahu / značného rozsahu*
- *S – praep S:* *na zdraví / na majetku / na pozemku*
- *S – VF:* *vznikla / přesahuje / hrozí*
- *VF – S:* *zavinil / způsobil / zabránil (škodu/škodě)*
- *VF – praep S:* *dochází (ke škodám) / odpovídá (za škodu)*
- *S – (praep) S:* *zjišťování / náhrada / úhrada (škody) / odpovědnost (za škodu)*
- *INF – S:* *zabránit (škodě) / nahradit (škodu)*
- *INF – praep S:* *odpovídat (za škodu)*

● Sémantické hodnocení vazeb:

- Ad A - S
- Věcná povaha škody
- Vyjádření původu a příčiny škody
- Kvantitativní hodnocení škody
- Společenské hodnocení škody

● Model morfologicko-syntagmatické struktury



Analýza právních textů

- **D. Výrazová segmentace textu**
- **E. Vztah mezi formální větou a textovou strukturací**
- **F. Výskyt vyhledávaných syntagmat v segmentech textu**
- **G. Distinktivní vlastnosti spojovacích prostředků (spojky a obdobné výrazy)**
- **H. Morfologie, gramatika, sémantika právních textů**
- **I. Struktura právních textů a jejich formální vztahy**
- **J. Využití:**
 - **1. Tvorba právních předpisů**
 - Formulace koncepčního záměru
 - Koncepce konkrétní legislativní úpravy
 - Zjišťování kontextových souvislostí
 - Právnělingvistická podpora legislativní činnosti
 - Vytváření úplných znění právních předpisů
 - Zpracování důvodových zpráv
 - Příprava legislativních výzkumů
 - **2. Tvorba právní terminologie**
 - **3. Tvorba oborových klasifikací a tezurů**
 - **4. Vyhledávání v oborových databázích**