



Dokumentační služba projektu MediGrid

Dokumentování sémantiky
lékařských dat

Projekt MediGrid

■ Účastníci

- FN Motol
- CESNET z.s.p.o.
- Masarykova nemocnice Ústí nad Labem

■ Cíl

- návrh, vývoj a pilotní implementace MediGridu - prostředí a modulárního systému aplikací pro distribuované zpracování datových a výpočetních úloh ve zdravotnictví s použitím technologie Gridu

Uspořádání MediGridu

■ Moduly

- Automatizované nástroje pro transformaci indikátorů
- Indikátory a třídy indikátorů
 - záznam, který někdo pořídil sám (nebo pomocí automatu) pro své budoucí použití nebo pro použití někoho jiného (Husserl)
- Relace mezi třídami indikátorů

■ Řadič

- Automatické řetězení modulů

■ Dokumentační služba

- Jednoznačný popis sémantiky lékařských dat, zejména modulů a indikátorů

Jednoznačný popis dat a vztahů

■ Klasifikační a terminologické systémy

- MeSH, ICD, SNOMED

- Vlastnosti

 - Hierarchické uspořádání

 - Převládající typy vztahů mezi pojmy – taxonomie (is_a), synonymie (see, see also), meronymie (part_of)

■ Ontologie

- UMLS, On9, OpenGALEN

- „Explicitní specifikace konceptualizace“

- Problémy

 - Vágnost pojmů

 - Inkonzistence vztahů

 - Binární charakter vztahů

 - Malé pokrytí

Příklad: UMLS

- UMLS (Unified Medical Language System)
 - soubor databází (UMLS Knowledge Sources) a softwarových nástrojů určených pro budování elektronických informačních systémů, které vytvářejí, zpracovávají, vyhledávají, integrují či shromažďují medicínské informace
 - MetaThesaurus - multilinguální databáze biomedicínských konceptů
 - Semantic Network - informace o sémantických typech/kategoriích konceptů Metatezauru a všech vztazích mezi těmito typy, na jejichž základě lze odvodit vztahy mezi koncepty
 - Specialist Lexicon

Příklad: BMI pomocí UMLS

- Koncepty: body height, body weight, body mass index
- Odpovídající sémantické typy: organism attribute, quantitative concept, clinical attribute
- Vztahy: is_a, associated_with, degree_of
 - Nepopisují dostatečně výpočet BMI

Řešení v DS MediGridu

- Ad-hoc (ex-post) ontologie
 - Generovaná automaticky na základě strojem zpracovatelných součástí DS

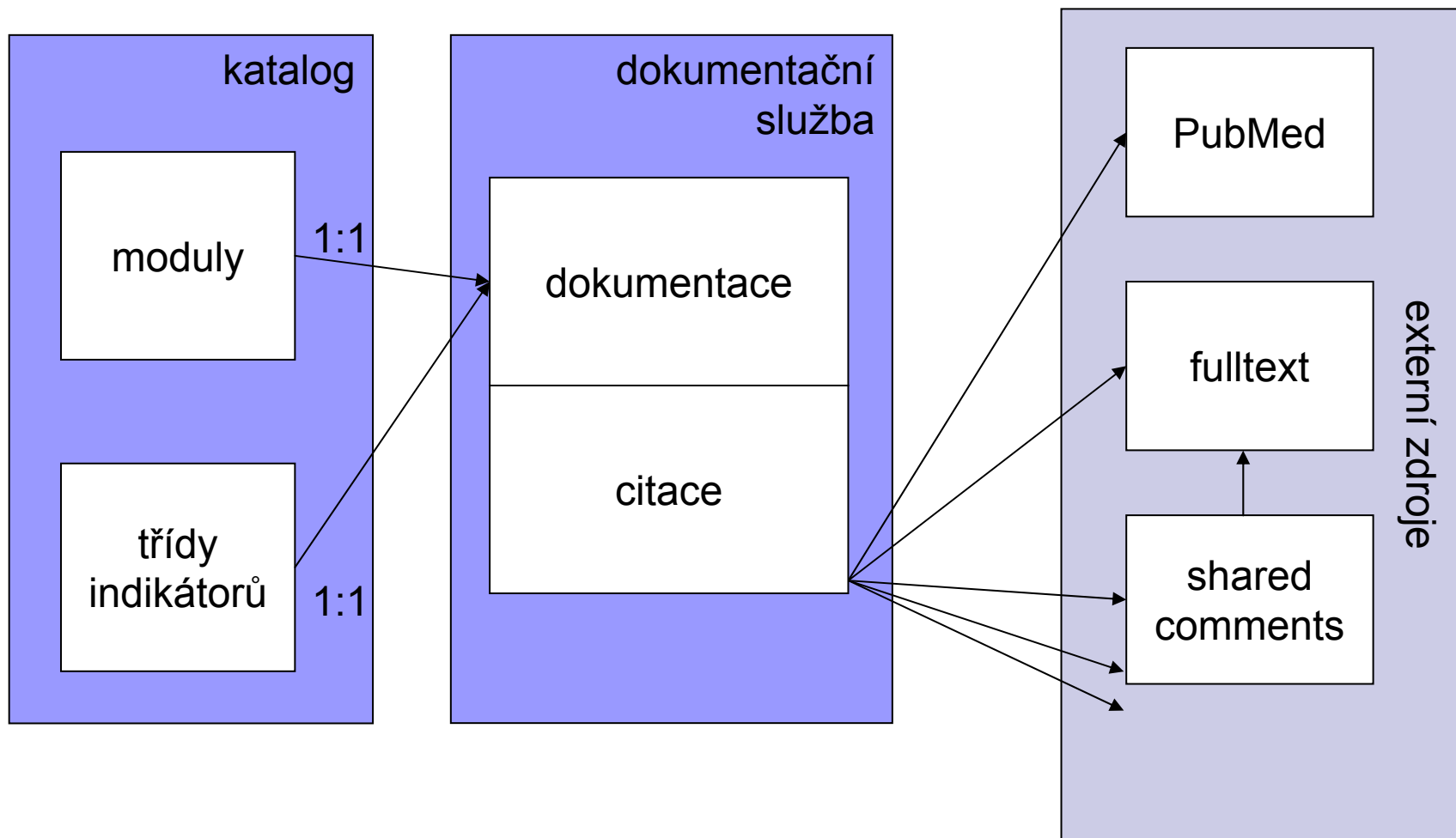
Dokumentační služba

- Jednoznačný popis sémantiky lékařských dat, zejména modulů a indikátorů
- Prostředek pro komunikaci odborníků
- Sdílení unikátních znalostí
- Podklad pro hodnocení v rámci EBM

Obsah DS

- Dokumentace
 - Modulů
 - Tříd indikátorů
- Citace (bibliografické záznamy)
- Strojem zpracovatelná podoba → ad-hoc ontologie
- Lidsky čitelný popis doplněný citacemi relevantních dokumentů

Schema dokumentační služby



Dokumentace tříd indikátorů a modulů

- Koncepty UMLS, případně jiných kontrolovaných slovníků nebo slovníku uživatelského
- Autor
- Popis v rozsahu nezbytném pro odlišení od jiných tříd indikátorů
- Relevantní citace

Struktura dokumentace

- zachycení odborného popisu
 - strukturovaný text
 - členění do funkčních bloků
- kontrola formálních náležitostí
- základní informace pro implementaci modulu

Ukládání citací

- Jednoduchý „reference manager“
 - uložení záznamu jakékoliv publikované i nepublikované informace, a to manuálně či stažením záznamu z externí databáze
 - vložení citaci do textu dokumentace
 - propojení citace na externí zdroj, např. fulltext či bibliografický záznam ze zdrojové databáze
 - Doplněkové funkce pro účely MediGridu
 - Hodnocení kvality citace – poloautomaticky na základě uživatelem zadaných údajů
 - Specifikace vztahu mezi citujícím a citovaným dokumentem

Struktura citace

■ Požadavky

- XML schéma
- jednoduše použitelná na aplikační úrovni
- Obsahově splňuje minimálně požadavky IJCME (vancouverská konvence)
- Libovolně rozšiřitelná

■ Volba formátu

- DC, MODS/MARC, BibTeXXML, DIVA, EndNote, Prism, PubMed, Sharef
- vlastní

Závěr

- Model DS bude ověřen v pilotním provozu
- Vše o projektu: <http://www.medigrid.cz>
- Kontakty
 - ajarolimkova@seznam.cz
 - petr.lesny@lfmotol.cuni.cz