

# MediGrid v1

*Petr Lesný, Kryštof Slabý, Tomáš Holeček,  
Adéla Jarolímková, Jan Vejvalka*

# Obsah

- Úvod (co, proč)
- Dokumentační služba
  - Účel
  - Struktura
  - Dokumentované entity
- Realizace
  - Klient, řadič a moduly
- Závěr
  - Nové otázky



# Úvod: Co?

- <http://medigrid.lf2.cuni.cz>
- Fungující systém, schopný řetězit odborné moduly a odborně dokumentovat entity, na které se odkazujeme
- Implementace jako WebServices (gSOAP)
- Klient na platformě .NET

# Úvod: Proč?

- Motivace spolupracujících odborníků
- Vznik odborné informační základny, ze které mohou čerpat týmy programátorů
- Implementační nenáročnost → snadná implementace v GRIDových centrech
- Prezentace na mezinárodních konferencích + publikace
- *Proof-of-concept* verze 0 pomůže položit si ty správné otázky pro verzi 1

# Proč dokumentujeme

- Abychom nezapomněli
- Komunikace autorů z jednotlivých odborných domén
- „Evidence-based“ medicína
  - Abychom byli schopni hodnotit v rámci EBM
- Postupný vývoj a upřesňování znalostí
- Unikátní zdroje znalostí
  - Obvykle jeden relevantní autor / tým
  - Doplnění, diskuse
  - Recenzenti, oponentury, ...

# Vrstvy dokumentace

## 1. Seznam zdrojů dostupných v systému

- Později inherentní prostředky sítě GRID

## 2. „Technická data“

- Typy dat / prezentace hodnot indikátorů
- Vstupní a výstupní indikátory modulů

~~✂ → AD-HOC ONTOLOGIE~~

## 3. Odborná dokumentace

- = vlastní dokumentační služba

# Dokumentační služba

- Základ jakéhokoliv systému, který vyžaduje spolupráci odborníků z různých oborů medicíny
- Ve verzi 0 jsou jednotlivé úrovně dokumentace sloučené
  - Snadný přístup týmů programátorů
  - Snadný strojový přístup

# Kdo dokumentuje

- Úkol programátorů a SWIŽ
  - Implementace nástrojů
  - Efektivita, distribuce znalostí v systému
- Úkol knihovníků a medicínských informatiků
  - Struktura dokumentace
  - Mechanismy, včetně kontroly kvality
  - Usnadnění spolupráce týmů
- Úkol specializovaných odborníků (lékařů)
  - Obsah dokumentační služby, údržba
  - Diskuse, oponentury, recenze

# Co dokumentujeme

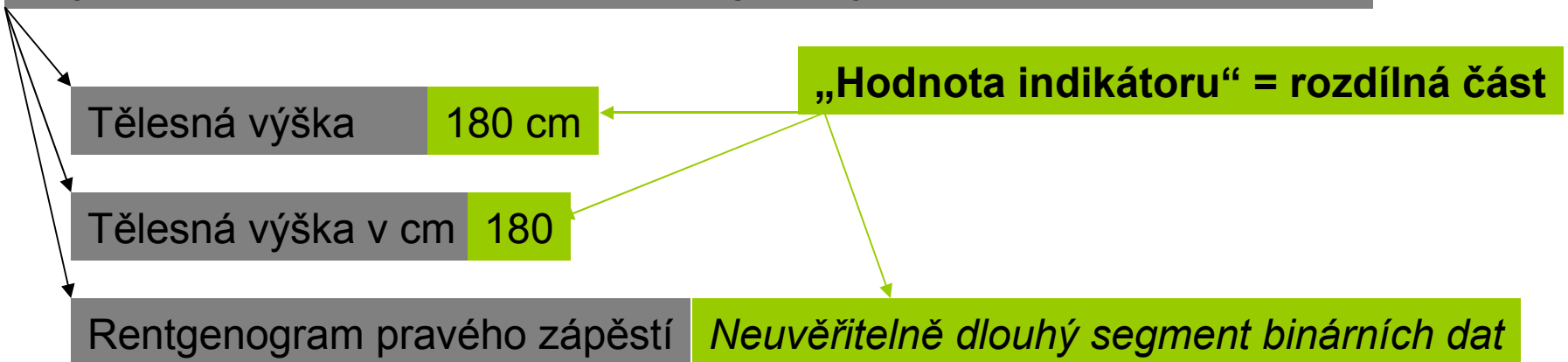
- **Třída indikátorů**
  - Sdružuje všechny indikátory, které patří ke stejnému pojmu
- **Relace mezi třídami indikátorů (=Modul)**
  - Implementována jako webová služba
- **Citace**
  - Odkazy do literatury, na které se odvoláváme při popisu tříd indikátorů i relací mezi nimi

# Třída indikátorů - terminologie

- Indikátorem něco sdělují (buď sám sobě, nebo jinému čtenáři)
- Záznam indikátorů ve verzi 0
  - Pracujeme s indikátory, které můžeme rozložit na dvě části:

## Identifikace třídy indikátorů

Pojem, resp. „primární termín“, který k pojmu patří = společná část



# Dokumentace třídy indikátorů

- Unikátní identifikátor
- Primární termín
- Popis
- Odkazy (→ **Citace, Moduly, jiné třídy indikátorů**)
- Vstupní a výstupní definice
  - Návod, jak může uživatel zadávat data, která patří do této třídy indikátorů (→ upřesnění)
  - Návod, jak data z této třídy indikátorů zobrazit na výstupu uživateli

# Dokumentace třídy indikátorů

- Popis
  - Nejdůležitější část dokumentace
  - Musí kompletně vysvětlovat význam dat, která jsou obsahem „upřesnění“ (těch dat, kterými se odlišují indikátory patřící do jedné třídy indikátorů)
    - Jednotky?
    - Vztah k jiným třídám indikátorů?
    - Formát dat?
  - **Odkazy na citace, moduly a jiné třídy indikátorů**

# Dokumentace třídy indikátorů

- Vstup
  - Text (jedno i víceřádkový)
  - Číslo (minimum, maximum, jednotka)
  - Výběr hodnoty ze seznamu
  - Není možné zadávat
- Výstup
  - Text (jedno i víceřádkový)
  - Číslo (počet desetinných míst, jednotka)
  - Obrázek (formát – zatím jen SVG)

# Dokumentace modulu

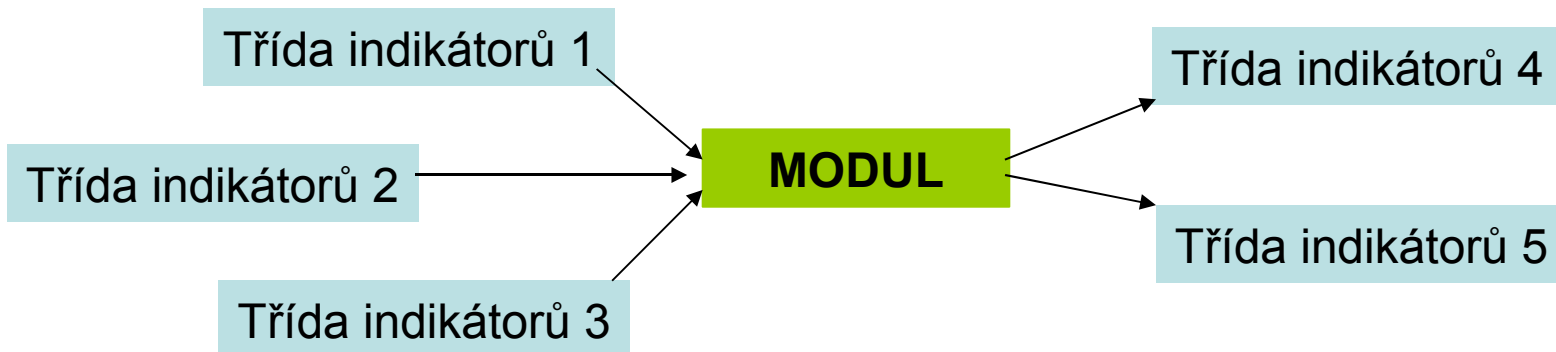
- Unikátní identifikátor
- Popis
- Odkazy na jiné moduly, citace, třídy indikátorů
- Seznam vstupních a výstupních tříd indikátorů, mezi kterými je relace definována
- Testovací hodnoty
  - Několik platných relací, které ověřují činnost modulu

# Dokumentace modulu

- Popis modulu = nejdůležitější část
  - Proč zrovna popsán algoritmus
  - Obsahuje nějaké sporné konstanty nebo konstanty, které se mohou později změnit?
  - Jsou nějaké alternativy?
  - Hraniční podmínky?
  - Byl ověřen v rámci Evidence based medicine (studie, jak)
  - Byl nějakými autory zpochybněn?

# Dokumentace modulu

- Popis modulu = nejdůležitější část
  - Musí zcela vysvětlovat činnost modulu
  - Není nutné uvádět detailně algoritmus
- Odkazy na citace, jiné moduly, nebo třídy indikátorů



# Dokumentace modulu

- Algoritmus
  - Textový popis (→ někdo to implementuje)
  - Pseudojazyk LUA *manuál přiložen*

```
h = get_indicator("lesny:height") / 100.0
w = get_indicator("lesny:weight")

bmi = w / (h^2)

if bmi > 0.0 and bmi < 50.0 then
    set_indicator("lesny:bmi", bmi)
end
```



Můžete hned  
otestovat

[reklama]

Podobný **Pascalu**; <http://www.lua.org>

Umí pracovat s čísly, řetězci a asociativními tabulkami

# Implementace modulu

- Implementace vlastní webové služby v C++
  - Účet na medigrid.lf2.cuni.cz
  - gSOAP-2.7 + jednoduchá makra + make...

```
#include "mdl.h"
BIND_PORT(9000)
INPUT
    double vyska_cm = DOUBLE_VALUE("lesny:height") / 100.0;
    double hmotnost_kg = DOUBLE_VALUE("lesny:weight");
ASSERT(vyska_cm > 0.0);
FUNCTION
    double bmi_kgm2 = hmotnost_kg / vyska_cm / vyska_cm;
OUTPUT
    SAVE_DOUBLE("lesny:bmi", bmi_kgm2);
END
```

– Pro ostatní jazyky je k dispozici soubor WSDL

# Dokumentace modulu

- Testy – později automatizované testování
  - Čím složitější algoritmus, tím více testů
- URI
  - Pro LUA moduly je generováno automaticky
  - Jinak

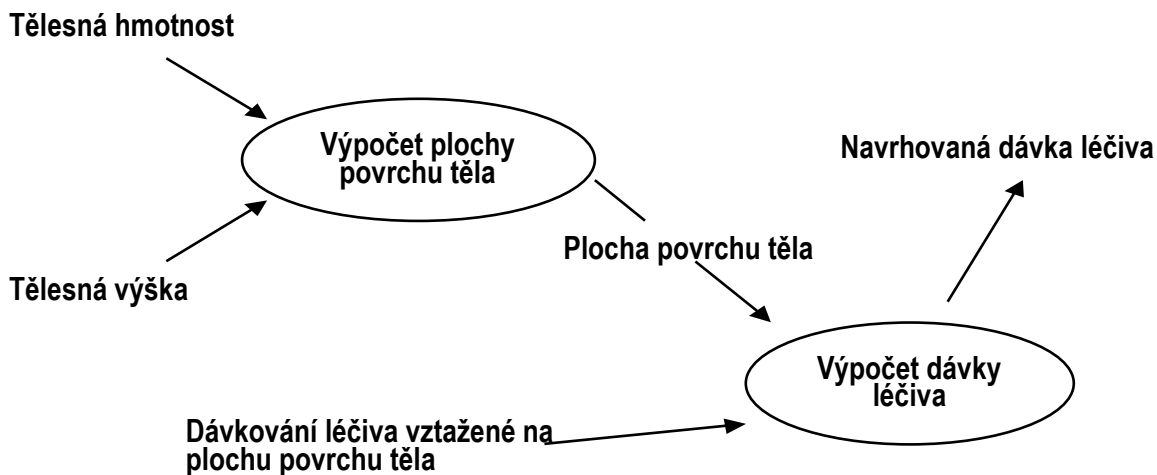
[http://medigrid.lf2.cuni.cz:přidělené\\_číslo](http://medigrid.lf2.cuni.cz:přidělené_číslo)

# Dokumentace citace

- Ve verzi 0 maximálně zjednodušená
- Unikátní identifikátor
- Citace ve Vancouverské notaci
  - EndNote ji umí generovat automaticky
  - detailní popis přiložen
- PMID
  - Pro články, které byly získány z PubMed

# Ad-hoc ontologie

- Relace mezi třídami indikátorů
- Generována z údajů, které jsou dostupné v dokumentační službě
- Přidání nového modulu (relace) → změna



# Klient DS

- Windows XP / 2000
- Napsán v C# .NET (zdrojové kódy k dispozici na stránkách <http://medigrid.lf2.cuni.cz/download>)
- Komentáře zasílejte na [petr.lesny@lfmotel.cuni.cz](mailto:petr.lesny@lfmotel.cuni.cz)

# Klient DS

The screenshot displays the MediGrid documentation service v0 client interface. The main window shows a citation for "Blaha, P. a kol.: V. celostatni antropologicky vyzkum deti a mladez...". The citation ID is "lesny:cav199".

The "Run module" window is open, showing the module name "Body mass index" and the URL "http://mediqid.lf2.cuni.cz:9000?lesny:bmi".

The "New module..." window is also open, showing the configuration for input and output indicator classes. The "Input indicator classes" and "Output indicator classes" sections are empty, with "Remove", "Add", and "Browse my..." buttons available for each.

The "List of all items" window shows a tree view of the database structure, including "lesny" and "citati".

The status bar at the bottom shows the URL "http://mediqid.lf2.cuni.cz:8100" and the message "Communication ok".

# Realizace

- V současnosti možnost implementovat libovolný algoritmus, který operuje v rámci jednoho kontextu
- Řadič: propojení algoritmů v rámci jednoho kontextu
- Vstup, výstup – DS poskytuje nástroje pro dynamickou konstrukci vstupně výstupních rozhraní

# Realizace

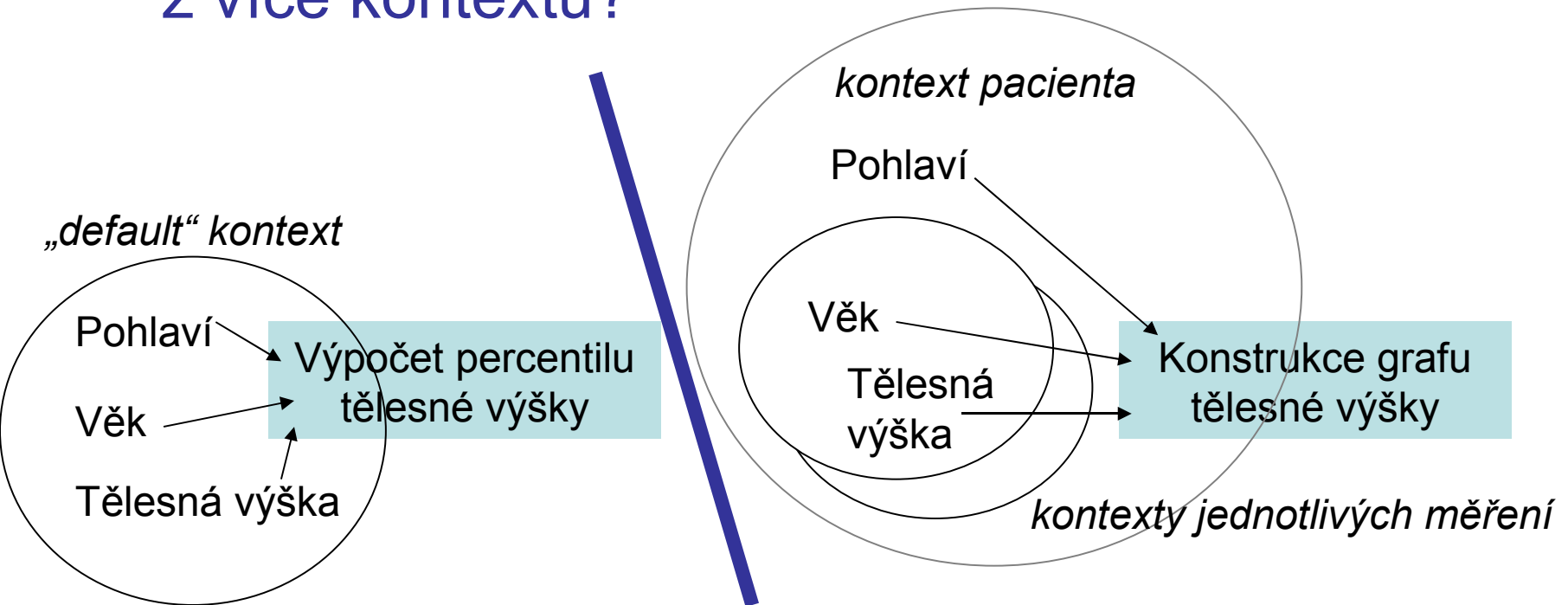
- Webové kalkulačky – MEDAL apod.
- Žurnál – záznam o proběhlém výpočtu (zřetězení relací), prostřednictvím kterého je možné výpočet kdykoliv zopakovat

# Otázky po verzi 0...

1. Kontexty
2. Struktura dokumentace a odkazů
3. Znalostní domény
4. Kontrola a správa verzí
5. Správa důvěry
6. Uložení dat
7. Implementace v GRIDovém prostředí

# Kontexty

- Relace ve verzi 0 jsou pouze nad daty v „default“ kontextu
- Jak definovat kontexty?
- Jak definovat relace, které vyžadují indikátory z více kontextů?



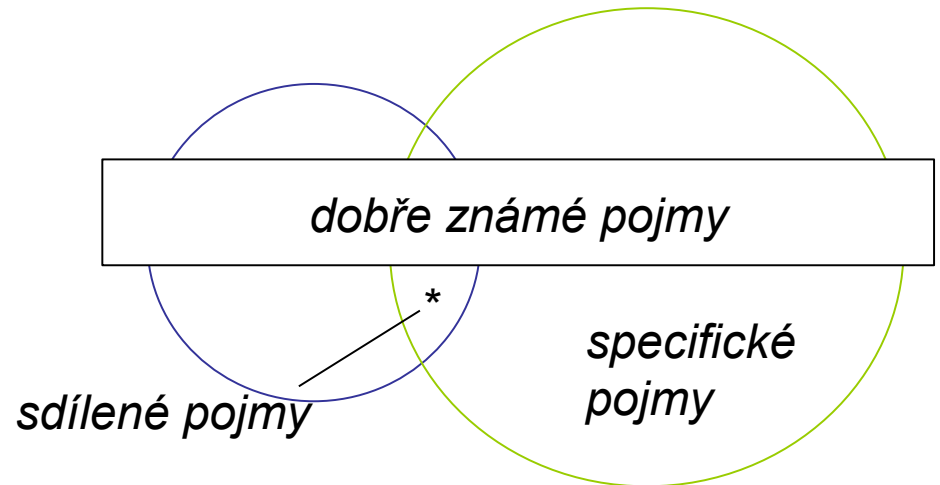
# Dokumentace a odkazy

- Jak udržovat dokumentaci?
- Jak se odkazovat do literatury nebo na jiné myslitelné zdroje informací?
- Zhodnocení kvality a relevance informací?

  Kryšot, Adéla

# Znalostní domény

- Spolupráce týmů odborníků
- Každý tým má svoji znalostní doménu
  - Specifické pojmy
  - Pojmy sdílené s jinými týmy
  - Dobře známé pojmy (slovník)
- Ve verzi 0 odráží pojmenování unikátních identifikátorů  
(`lesny:percentil-BMI`)



# Verze

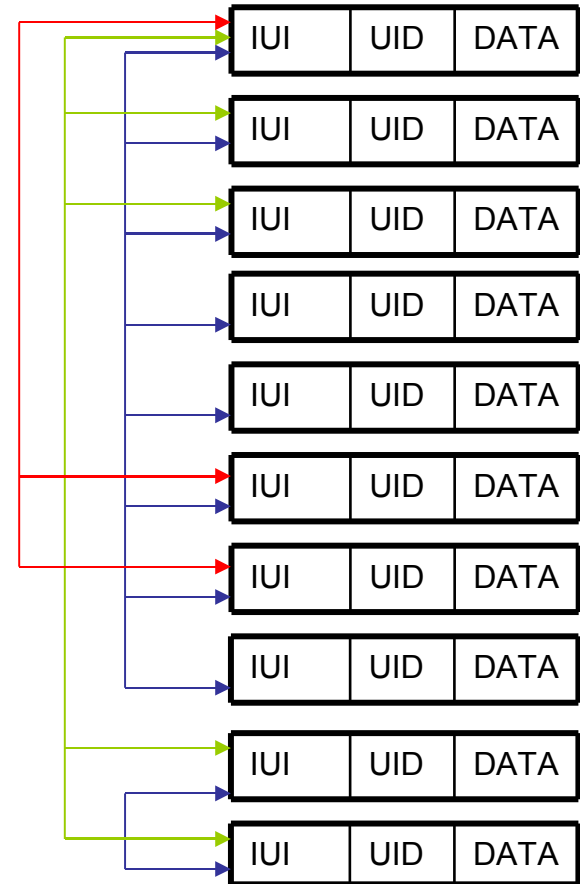
- Potřebujeme je?
  - - lékařské algoritmy nemají čísla verzí, odkazy probíhají přes citace (tedy DS)
  - + implementace má čísla verzí
  - - algoritmy buď platí nebo byly vyvráceny
  - + reprodukovatelnost výsledků
- Evidence based medicine
  - Rozhodování na podkladě dostupné evidence
  - Bez verzí – vždy poslední znalosti

# Správa důvěry

- Odborná záležitost
- Kontrola kvality
- MediGrid Reviewers' board ??
- Ve verzi 0 řešeno jednoduchým seznamem autorit, které dané relaci důvěřují

# Uložení dat

- Indikátory seskupené do kontextů
  - Generický formát pro uložení biomedicínských dat?
- **Není relační databáze**
  - Může být relační databází simulováno
  - IUI



# Implementace v GRIDu

- V tomto stádiu prospěšná – pomůže strukturovat DS
- Podpořit vznik GRIDových center, ve kterých budou umístěny moduly
  - Praha, Brno, Ústí nad Labem,...
- Kvalitní zpracování
- Doplnit OGSA-compliance pro stávající implementaci
- Přenést části DS do UDDI (?jakou?)

Děkuji za pozornost  
a vyzývám k diskuzi