

STRUČNÉ PŘEDSTAVENÍ SNOMED CT

Libor Seidl

Anotace

Přehledový článek seznámí v přiměřeném detailu s řízenou klinickou terminologií SNOMED CT, se souvisejícími přístupy, postupy, principy a technologiemi. Hlavními tématy jsou terminologický obsah, logický model, principy a možnosti pre- a post-koordinace a využití SNOMED CT v grafickém uživatelském rozhraní klinického informačního systému. Článek odkazuje na klíčové zdroje o SNOMED CT a tak může posloužit jako výchozí studijní materiál pro hlubší seznámení se a detailní studium SNOMED CT.

Klíčová slova

klinická terminologie, SNOMED-CT, logický model, post-koordinace, grafické uživatelské rozhraní (GUI)

1 Úvod

Zavedení řízené klinické terminologie SNOMED CT do praxe v českých podmínkách je jedním z vytyčených cílů nově schválené Národní strategie elektronického zdravotnictví[1]. Účelem tohoto přehledového článku je tedy předložit čtenáři přiměřeně detailní vzhled na SNOMED CT a související přístupy, postupy a technologie otevírající nové možnosti při sběru, uložení a zpracování klinických informací o pacientovi. Přestože momentálně jsme v České republice velmi vzdáleni českému překladu SNOMED CT, tak přístupy, implementační postupy i benefity jsou zřetelné i při použití anglické verze. Navíc následný přechod na českou verzi je spíše formální, než technologicky složitý krok.

Česká republika je prostřednictvím KSRZIS a nyní UZIS členem sdružení SNOMED International (původně IHTSDO – International Health Terminology Standards Organisation). Toto členství přináší možnost používat SNOMED CT na území České republiky prakticky bez omezení. Z pohledu autorského práva je ovšem potřeba SNOMED CT považovat za databázové dílo, které všeobecně požívá nejvyšší autorské ochrany a smí být použito pouze vyjmenovaným způsobem¹. Provádění a vydávání národního překladu je tak vyhrazeno pouze státem určené autoritě (tzv. National Release Center), firemní či lokální rozšíření je možné realizovat pouze v samostatných (a předem zaregistrovaných) modulech, použití mimo členské země podléhá registraci a poplatku, a doporučuje se registrovat každou organizaci užívající SNOMED CT v systému MLDS pro automatizované získávání pravidelných aktualizací. Aktuální verzi SNOMED CT je možné získat po registraci na <https://mlds.ihtsdotools.org/#/landing/CZ?lang=en>.

2 Terminologický obsah SNOMED CT

SNOMED CT je kontrolovanou klinickou terminologií, tedy její obsah je centrálně řízen v podobě schvalovacího procesu a následného přidělení unikátního identifikátoru

¹ [2], konkrétně § 90 – Obsah zvláštního práva pořizovatele databáze a § 91 – Omezení zvláštního práva pořizovatele databáze.

SCTID pro každý nový koncept. Termíny v jazykových mutacích jsou pak kontrolovány jednotlivými NRC.

Lednová edice 2017 SNOMED CT obsahuje celkem 326734 aktivních konceptů, okolo 1,2 milionu anglických popisků (Terms či Descriptions) a přes 700 000 definovaných vztahů mezi koncepty. Všechny koncepty SNOMED CT jsou uspořádány do celkem 19ti větví, kde vztah podřazenosti je ve významu vyššího detailu a užšího významu (anglicky „is a kind of“). Obsah každé větve tvoří striktní klasifikaci, takže koncepty umístěné hlouběji se typicky vyskytují pod několika obecnějšími koncepty. Tato struktura umožňuje uživatelům více přístupů či pohledů ke stejné problematice, což je typické pro způsob vyjadřování i uvažování, avšak znesnadňuje to statistické zpracování (např. ve srovnání s ICD10).

Jedenáct větví SNOMED CT je čistě klinických:

- Body structure (body structure),
- Clinical finding (finding),
- Observable entity (observable entity),
- Organism (organism),
- Pharmaceutical / biologic product (product),
- Procedure (procedure),
- Qualifier value (qualifier value),
- Situation with explicit context (situation),
- Specimen (specimen),
- Stages and scales (staging scale),
- Substance (substance),

Pět větví utváří kontext pacienta i celého výrazu:

- Environment or geographical location (environment / location),
- Event (event),
- Physical force (physical force),
- Physical object (physical object),
- Social context (social concept),

A dále tři větve pracují s doménou informačních technologií a propojují ji s klinickým obsahem:

- Record artifact (record artifact),
- SNOMED CT Model Component (metadata),
- Special concept (special concept).

Vzhledem k neexistenci českého překladu uvádím anglické názvy základních větví.

V závorce je vždy uveden tzv. sémantický tag uváděný u všech konceptů z dané větve.

Zatímco zejména klinické větve jsou předmětem intenzivního výzkumu, třídění a re-strukturování, jiné větve (např. „Physical object“) sdružují pouze koncepty nutné např. pro mapování na ICD10 a nejsou výrazně strukturovány. Obsah SNOMED CT je možné procházet v SNOMED CT Browsersu [3], případně i v jiných nástrojích (ART-DECOR, Web NLM aj.)

3 Logický model

Logický model SNOMED CT představuje datovou strukturu umožňující popis celého obsahu SNOMED CT. Vedle datového modelu se zároveň jedná o definici datového formátu pro vydávání SNOMED CT (tzv. Release Format ver. 2) a taktéž se jedná o jednu z možných datových struktur pro uchování SNOMED CT v relační databázi.

Logický model SNOMED CT je vyjádřen na obrázku 1 jako class diagram. Ústřední třídou je *Component* nesoucí identifikátor (SCTID), datum počátku platnosti a příznak zneplatnění. Hlavními komponentami SNOMED CT jsou koncept reprezentující klinickou myšlenku, *popisek* (Description) přiřazený ke konceptu, kterému udává název a popis; a *jednosměrná vazba* (Relationship) propojující dva koncepty vazbou typu „is_a“, případně jiným typem.

Aby se logický model vyhnul různým číselníkům definujícím sémantiku modelu (např. typy vazeb, typy popisů, nebo definiční status konceptu), používá logický model u těchto atributů datový typ SCTID a referenci na existující koncepty definované ve SNOMED CT ve větvi „Model Component“. Národní jazyková verze tak automaticky poskytuje i překlad pro jednotlivé hodnoty atributů a tedy i pro celou sémantiku logického modelu.

Referenční množiny (Refsety) jsou zvláštní třídou logického modelu, která umožňuje definovat paralelní množiny k základní hierarchii 19 větví zmíněných výše. Prostá

referenční množina (Simple Refset) je vyjmenování konkrétních komponent (tj. nejen konceptů), které dohromady tvoří smysluplný celek. Typickým příkladem je seskupení všech konceptů pro seznam alergenů. S tímto seznamem pak umí informační systémy lehce pracovat. Jiným příkladem je seskupení všech *Descriptions* pro konkrétní jazyk nebo dialekt.

Refsety různých typů mají své vlastní datové modely, aby mohly nést dodatečné informace pro konkrétní účel. Například Refset pro mapování SNOMED CT do/z externího číselníku potřebuje uvést kód z tohoto číselníku (tedy pravděpodobně jako String).

4 Pre / Post-koordinace

Každý koncept ve SNOMED CT reprezentuje unikátní klinickou myšlenku [5]. Přestože popis konceptu (název) spolu se sémantickým tagem je pravděpodobně pro klinika jednoznačným identifikátorem takové myšlenky, počítač tomuto popisku nerozumí a musí se opírat o formální popis konceptu pomocí zatřídění vazbami „is_a“ k nadřazeným konceptům a definičními vazbami na další koncepty.

Pokud je klinická myšlenka přímo zastoupena ve SNOMED CT konceptem, říkáme, že koncept je pre-koordinovaný. Pokud je zároveň přítomna úplná definice pomocí vazeb na další koncepty, pak je tento koncept označen za úplně definovaný (Fully Defined Concept).

Vyjmenování všech klinických myšlenek ve SNOMED CT by vedlo v kombinatorickou explozi, nadměrný počet konceptů a praktickou nepoužitelnost celé terminologie. Navíc mnoho kombinací je nesmyslných, a mnoho funkčních kombinací stejně nemá v klinické praxi výstižné jméno a jsou vždy opisovány větou vedlejší, ukazovacím zájmenem či jinak. Z toho důvodu je i ve SNOMED CT zavedena gramatika (Compositional Grammar)[6]. Výrazem podle této gramatiky lze doplnit (resp. dodefinovat) již stávající koncept o nové detaily. Taková nová klinická myšlenka (např. doplnění levo/pravostrannosti orgánu) se označuje jako *post-koordinovaný koncept*.

Pokud se náhodou v některé z následujících verzí SNOMED CT objeví tato klinická myšlenka jako plně definovaný pre-koordinovaný koncept, lze následně spočítat, zda dříve post-koordinovaný koncept není ekvivalentní s touto novou definicí. Historická data založená na post-koordinaci tak mohou být udržována v homogenním stavu i k novějším verzím SNOMED CT.

Příklad post-koordinace pomocí Compositional grammar, převzato z [7]:

```
=== 56265001 |Heart disease (disorder)| :
{ 363698007 |Finding site (attribute)|
= 58095006 |Interatrial septum structure (body structure)|,
116676008 |Associated morphology (attribute)|
= 49755003 |Morphologically abnormal structure (morphologic
abnormality)|
}
```

5 Použití SNOMED CT v grafickém uživatelském rozhraní

Jednotlivé popisky (Descriptions) pre-koordinovaných konceptů ve SNOMED CT mohou být přímo použity v grafickém uživatelském rozhraní (GUI). Výhodou je nejen případný automatický překlad zobrazovaných dat do jazyka uživatele, či apriorní volba správných pojmů z terminologie, ale i možnost využít oficiálních či vlastních refsetů pro omezení nebo zkrácení výběrových seznamů (tzv. listboxů, comboboxů, roletových menu atp.). Tyto seznamy mohou být navíc filtrovány podle klinické relevance odvozené z kontextu aktuálního zobrazení a tak ještě více usnadňovat a zefektivňovat práci klinika s počítačem.

Novou a zajímavou možností je využití principu postkoordinace konceptu v GUI a následného přepočtu na případnou pre-koordinovanou hodnotu. Na následujícím obrázku jsou vyobrazeny tři přístupy pro zadání diagnózy, všechny vedou na stejnou hodnotu uloženou do databáze. Zvýšený komfort pro uživatele je zde zcela neoddiskutovatelný.

6 Přizpůsobení terminologie lokálnímu použití

Jak bylo již výše zmíněno, popisky (Descriptions) k jednotlivým konceptům jsou seskupeny do jazykového RefSetu. Volbou tohoto refsetu lze prakticky za chodu přepínat mezi jazyky, kterými jsou data ve SNOMED CT reprezentována.

Za splnění určitých administrativních náležitostí je ale možné vytvořit si vlastní jazykový refset obsahující zkratky, preferované slangové výrazy a jiné jazykové zvláštnosti konkrétního kraje, konkrétního zařízení, nebo dokonce konkrétního nemocničního oddělení či uživatele. Jedna datová věta z informačního systému tak může být i v rámci tohoto informačního systému zobrazena dvěma uživatelům odlišně, respektující jejich preferovanou sadu zkratk a slangových výrazů. To samozřejmě platí i pro veškeré ostatní GUI, včetně vkládání nových klinických údajů a rozpoznávání volného textu.

Tento přístup silně popírá dnešní zavedenou praxi při certifikaci nemocnic, kdy se

požaduje unifikovaný seznam užívaných zkratk ve zdravotnické dokumentaci[8], který je následně závazný pro veškerý personál. Zvýšený komfort pro uživatele je zde opět neoddiskutovatelný.

Závěr

Řízená klinická terminologie může přispět k rozvoji strukturovaného klinického záznamu v informačních systémech. To umožní automatizaci některých administrativních procesů, automatizaci statistického zjišťování nebo automatizaci vykazování zdravotním pojišťovnám. Mezi významné přínosy patří také efekty plynoucí z využití SNOMED CT v grafickém uživatelském rozhraní, kde umožňuje personalizaci zobrazení, i zadávání dat rozličným způsobem, avšak při udržení konzistence výsledných hodnot uložených v databázi.

Aby se SNOMED CT dostal do běžné praxe, bude nutné nejen zajistit prvotní překlad do češtiny a následné průběžné rozšiřování a udržování (úloha pro National release center), ale také inovace další infrastruktury vně poskytovatelů zdravotní péče nezbytné pro podporu výpočetních úkonů nad strukturovanou zdravotnickou dokumentací. Mezi tyto inovace bezesporu patří:

- formální popis zákonných požadavků na obsah zdravotnické dokumentace,
- algoritmická definice položek statistického zjišťování ze strany UZIS,
- algoritmická definice položek sazebníku zdravotní pojišťovny (mapování mezi povinným obsahem zdravotnické dokumentace a konkrétní položkou sazebníku),

a v neposlední řadě také algoritmizace klinických doporučených postupů a alertů na straně odborných společností.

Literatura

- [1.] MINISTERSTVO ZDRAVOTNICTVÍ ČESKÉ REPUBLIKY. *Národní strategie elektronického zdravotnictví*. Verze 1.00. Praha, 2016. Licencováno pod CC BY 4.0, [cit. 2017-02-23]. Dostupné z: http://www.nsez.cz/dokumenty/schvalena-narodni-strategie-elektronickeho-zdravotnictvi-na-obdobi-2016-2020_13051_3558_31.html
- [2.] Zákon č. 121/2000 Sb., o právu autorském, o právech souvisejících s právem autorským a o změně některých zákonů (autorský zákon)
- [3.] SNOMED International. *SNOMED CT Browser*, [online], Dostupné z: <http://browser.ihtsdotools.org/>
- [4.] ELC: IHTSDO. *SNOMED CT E-Learning Center*, [online], Dostupné z: <http://doc.ihtsdo.org/elearning.html>
- [5.] IHTSDO. *SNOMED CT Starter Guide*, 2016 [online] [cit. 2016-09-22], Dostupné z http://snomed.org/sq_gb.pdf.html
- [6.] IHTSDO. *SNOMED CT Technical Implementation Guide*, 2016 [online] [cit. 2016-09-22], Dostupné z: http://snomed.org/tig_gb.pdf.html
- [7.] Wikipedia contributors. *SNOMED CT / Precoordination and Postcoordination*, Wikipedia, *The Free Encyclopedia* [online] [cit. 2016-09-22], Dostupné z: https://en.wikipedia.org/wiki/SNOMED_CT#Precoordination_and_postcoordination.
- [8.] *Spojená akreditační komise o.p.s. Akreditační standardy pro nemocnice*, 3. vydání, 2013 [online], ISBN 978-80-87323-04-05, Dostupné z: http://www.sakcr.cz/get_file.php?file=/admin/upload/download/sak-standardy-nemocnice-2014.pdf

Kontakt

Libor Seidl

Ministerstvo zdravotnictví České republiky

Palackého nám. 4

128 01 Praha 2

tel: +420 224 97 2608

e-mail: libor.seidl@mzcr.cz