

**Posudek oponenta disertační práce PhDr. Kláry Rösslerové  
„Historie, současnost a budoucnost výměnných formátů  
bibliografických dat“  
předkládané v roce 2016  
na Ústavu informačních studií a knihovnictví FF UK**

## **1 Stručná charakteristika práce**

Práce se věnuje tématu, které je nejen zajímavé, ale je současně velmi aktuální. Navíc je nutno ještě zmínit i případnou praktickou použitelnost. Proto je vznik podobné práce velmi záslužný.

Práce je rozdělena do deseti kapitol. Kapitoly 2–3 vymezují nezbytnou terminologii a typologii s ohledem na další text práce, následuje kapitola 4 věnovaná podrobné analýze výměnných formátů pro bibliografická data. V kapitole 5 je provedena rekapitulace vývoje zavádění a používání výměnných formátů v České republice. Kapitola 6 charakterizuje nejpoužívanější formáty v celosvětovém měřítku. Kapitoly 7–8 jsou zaměřeny na novější trend publikování propojených dat. Kapitola 9 popisuje provedený experiment. Kapitola 10 analyzuje připravenost knihovních softwarů s ohledem na očekávaný budoucí vývoj výměnných formátů. Hlavní text završuje kapitola 11 věnovaná diskusi k postojům rozhodujících orgánů a institucí k budoucnosti výměnných formátů.

## **2 Stručné celkové zhodnocení práce**

Předložená dizertační práce je obsáhlá, ač není takto explicitně rozdělena, lze ji rozčlenit na dvě obsahové části. První část tvoří kapitoly 4–6 věnované historickému přehledu vývoje výměnných formátů jak ve světě, tak v českém prostředí. Této části nelze vytknout snad vůbec nic. Je zpracována se vši důkladností a přesností, a to včetně bohatých odkazů na další prameny. Kdyby se někdy uvažovalo o monografii zaměřené na výměnné formáty, pak texty zmíněných kapitol by mohly být převzaty takřka bez úprav.

Pomyslná druhá část začíná sedmou kapitolou a autorka se snaží odhadnout budoucí vývoj ohledně strukturování bibliografických dat. Trend publikování dat na webu v podobě otevřených propojených dat je nezpochybnitelný pro všechny instituce produkující nějaká data, tedy i knihovny nevyjímaje. Úvaha o používání otevřených propojených dat ve funkci výměnného formátu je oprávněná a má dobrou šanci na úspěch. Toto i lze vyvodit ze šetření provedených autorkou v různých odborných komunitách. Ne tak jednoznačně však dopadá úvaha o používání propojených otevřených dat v roli interního formátu knihovních systémů.

## 3 Podrobné zhodnocení práce

### 3.1 Struktura argumentace

Vzhledem k rozsahu práce se autorka nevyvarovala některých nepřesností:

- Na str. 28 se jako princip propojených dat uvádí „*podmět – predikát – předmět*“, v české terminologii bývá však obvyklejší uvádět „subjekt-predikát-objekt“.
- Na str. 31 se slibuje vysvětlení pojmu „komunikativní formát“, v dalším textu kap. 2 se to však neobjevuje.
- Tvrzení na str. 34, že „*Zcela samostatná struktura interního formátu a jeho nezávislost na formátech vstupních a výstupních systémovým knihovníkům, či spíše programátorům, velmi ulehčuje práci.*“ je silně zjednodušující. I při použití vlastního formátu je nutno brát do úvahy komplexnost a granularitu údajů, protože jinak nebude možné data snadno převádět do formátů jiných.
- Znakovou sadou se obvykle myslí něco jiného, než co je uvedeno na str. 34: „*Znaková sada je tvořena kódy polí, podpolí a identifikátory.*“.
- Na str. 91 autorka tvrdí, že systém CDS/ISIS neumožňoval opakování polí, ale současně pak na str. 94 píše o opakovatelném poli 100. Není to rozpor?
- Nejsm si jist, že obrázek 16 na str. 115 je dobrou ilustrací principu propojených dat. V tomto případě se jedná spíše o extrakci údajů z Wikipedie.
- Ze 14 stran kapitoly 9 je plných 50 % rozsahu tvořeno generovanými zdrojovými kódy. Navíc jsou ještě v textu vloženy dva půlstránkové náhledy obrazovek.
- Nechybí v tabulce 10 jeden sloupec? Komentář na str. 163 „*souhrnná tabulka systémů s uvedením zjištěné informace o poskytovaných exportních formátech (viz Tabulka 10), o tom, jestli firmy (producenti) na svých webech nebo v systémové dokumentaci uvedli možnost publikování v podobě propojených dat. V posledním sloupci je pak uvedeno, zda firma publikovala jakoukoliv informaci o otevřených datech na webu, v blogu či jinde.*“ není dle mého názoru v souladu s tím, že tabulka obsahuje dva sloupce zaměřené na formáty.
- Pokud je zjištěno, že 16 % nepodporuje a současně u 53 % nešlo zjistit, pak to nelze zobecnit do tvrzení, že 69 % nepracuje s otevřenými daty – viz str. 166. Podobně se argumentuje (bohužel) i u dalších otázkách.
- V práci se operuje s pojmem „*ontologie*“, na rozdíl od jiných však nebyl zaveden. Dále pak se pro bibliografická data uvádí ontologie BIBO, nejsou zmíněny žádné jiné alternativy.

- Na str. 176 je uveden seznam otázek použitých v jednom ze tří šetření. Pro otázku „*Bude jedna vedoucí struktura propojených dat?*“ je však uveden anglický překlad „*Will there be one leading data structure?*“, který neodpovídá. Na místě je proto otázka, jaké znění otázky respondenti opravdu dostali?

## 3.2 Formální úroveň

### 3.2.1 Úprava práce

Dizertační práce je výsledkem dlouholetého úsilí, a proto je vhodné věnovat trochu pozornosti i finálnímu vzhledu a úpravě práce. V tomto smyslu lze nalézt v práci rezervy.

- Obsah práce má celkem 7 stran, což je dáno tím, že jsou do něj zahrnuty všechny úrovně číslovaných nadpisů. Pro přehled by však zcela postačovalo uvést jen první tři úrovně.
- Počet šesti číslovaných úrovní nadpisů je diskutabilní, protože to někdy vede až k absurditám, že pod nadpisem je jediná věta.
- V celé práci je použito patkové písmo, proto působí trochu nelogicky, že Seznam zkratek na str. 25–26 je vyhotoven za pomoci písma bezpatkového.
- Proč jsou nadpisy druhé úrovně na stránce umísťovány s předsazením doleva? U žádné jiné úrovně nadpisů není předsazení použito.
- Podle použitého číslování stran má práce bez příloh celkem 189 stran, ale tento údaj je zkreslen tím, že úvodní formální výbava má 26 stran. Bylo by přehlednější, kdyby byla pro úvodní formality použita samostatná řada s použitím třeba římských číslic a Úvod pak začínal na čísle 1.
- Úprava na str. 95 je méně přehledná, kombinace výčtových bodů a samostatných textových odstavců není šťastná.
- Na str. 96 je text nezarovnaný do bloku.
- V některých místech se nadpisy téměř dotýkají předchozího obsahu – viz třeba str. 166, 167, 174... – je to matoucí.
- Vzhledu práce by prospělo použití dělení slov na konci řádků, v řadě míst vede zarovnávání do bloku trochu necitlivému „ředění“ textu.

### 3.2.2 Jazykové připomínky

Práce je napsána poměrně pečlivě, jazykových prohřešků lze najít jen málo:

- Na str. 32 ve vložených větách oddělených pomlčkami nejsou na místě velká počáteční písmena.

- Na str. 37 je poněkud strojově působící věta „*Tyto typy formátů však míchají typy formátů dle různých typologií.*“. Na str. 38 pak něco podobného: „*katalogizátor katalogizuje záznam v souladu s používanými katalogizačními pravidly.*“ – nehledě na to, že v tomto případě je formulace i lehce nepřesná, protože záznam vzniká jako produkt katalogizace určitého objektu (třeba knihy).
- Na počátku str. 41 je „*Na příklad*“ – mělo by se psát dohromady.
- V textu je opakovaně komoleno jméno zakladatele jak webu samotného, tak webu sémantického – správně je *Tim Berners-Lee*, nikoliv *Barners-Lee* (viz třeba str. 113).
- V textu je hojně používán pojem „*převaděč*“, pro software daného typu je však obvyklejší a srozumitelnější označení „*konvertor*“.
- Formulace typu „*... je znázorněn viz Diagram...*“ jsou nepřirozené.

### 3.3 Práce s prameny

Práce s použitými prameny je příkladná. Již jen fakt, že seznam pramenů zahrnuje celkem 165 položek, dokazuje dokonalou autorčinu rešersní přípravu. Rovněž tak lze vyzdvihnout provázanost textu a pramenů díky nesčetným citacím.

### 3.4 Vlastní přínos

Bývá zvykem, že v úvodu dizertační práce jsou uvedeny cíle práce. V tomto směru je předložená práce je trochu tajuplná. Lze tedy vyjít z úvní věty Předmluvy „*Cílem předkládané disertace je nastínění možného budoucího vývoje na poli výměnných formátů bibliografických dat a doporučení pro knihovny, jakým způsobem se změně přizpůsobit.*“ a dále pak ještě předposlední odstavec Úvodu „*Hlavními výstupy práce je tedy zjištění nejpoužívanějšího výměnného formátu, zhodnocení nutnosti jeho nahrazení jiným formátem, nastínění možného vývoje, zjištění připravenosti knihoven na případné změny a popsání jednotlivých kroků v rámci publikování bibliografických záznamů v podobě propojených dat včetně praktických příkladů.*“. V Závěru, kde by to bylo na správném místě, není vlastní přínos vyjádřen. Stručné zhodnocení nosných kapitol:

- Kapitoly 4 a 5 opravdu dobře a výstižně opravňují, aby v názvu práce byl zmíněn pojem *historie*. Obě kapitoly mají v souhrnu více než 60 stran vyčerpávajícího přehledu, jak šla historie ve světě i u nás, je uvedena spousta příkladů.
- Kapitola 9 je nejen kvůli nevyváženosti obsahu (viz výše), ale současně se mi nezdá vhodné zařazovat do dizertační práce popis práce při vyplňování webového formuláře. Navíc nazývat toto *experimentem* je trochu nadnesené.

- Kapitola 11 je poslední kapitolou před Závěrem, a proto lze usuzovat, že by měla jednak práci završovat, jednak do práce vetknout úvahu o budoucnosti výměnných formátů. Kapitola je pojata jako průzkum, když respondenty byli členové Stálého výboru Sekce katalogizace IFLA, členové Poradního výboru MARC a dále pak také odběratelé e-mailové konference BIBFRAME. Této kapitole však lze vytknout metodické chyby, plynoucí zejména z toho, že pro malé počty respondentů není moc vhodné používat čistě kvantitativní metody. V situaci, kdy má jedna odpověď váhu 20 % by to mělo být zřejmé.

## 4 Dotazy k obhajobě

Autorka by mohla v rámci obhajoby rozvést myšlenku o výhodách a nevýhodách, když by se formát propojených otevřených dat stal univerzálním formátem pro celý „životní cyklus“ dat v rámci knihovního systému.

## 5 Závěr

I přes výše uvedené připomínky se domnívám, že předložená dizertační práce obsahuje všechny náležitosti a splňuje požadavky na dizertační práce kladené, a proto **práci doporučuji k obhajobě**.

V Praze dne 17. ledna 2017

Doc. Ing. Vilém Sklenák, CSc.  
Katedra informačního  
a znalostního inženýrství  
Fakulta informatiky a statistiky  
Vysoká škola ekonomická v Praze