

Nové trendy v oblasti tvorby a prezentácie bibliografických dát

Mgr. Ján Grman, PhD.
SVOP spol. s r.o.



Prolog

- 40 / 20 rokov
- štandardy, konverzie, spôsoby a účely evidencie
- informatik s prvkami knihovníka

Bibliografické metadáta

Na čo?
Pre koho?
Ako?

Vývoj

- Od lístkov cez tabuľkové zoznamy až po štrukturované databázové záznamy
- Od lineárneho popisu „položiek“ cez tvorbu heslárov, externých entít až k tvorbe multimedialných a sémantických sietí
- Vývoj štandardov pre prenos, kódovanie, syntax zápisu a interpretáciou sémantických vzťahov

Využívame naplno čo vývoj priniesol ?

Nevyužité príležitosti

- Všetko je na Google
 - Nie všetko sa dá hľadáním typu Google nájsť
- Dajme mi plný text a pohnem zemou
 - Plný text je skvelý doplnok hľadania a skvelý pre rýchly prístup k obsahu (ak je zverejnený), ale pri hľadaní znižuje relevanciu výsledkov a zvyšuje informačné prehustenie
- Stroje všetko vyriešia – ľudí netreba
 - Mravenčiu prácu informačného profesionála nenahradí zatiaľ nič

Problémy

- Technologické zaostávanie systémov
 - téma nie len o peniazoch
- Priveľký dôraz na formu na úkor obsahu
 - čiarky v ISBD verzus kľúčové slová
- Vyhľadateľnosť dát
 - deskriptívne metadáta – pridaná hodnota
- Presnosť evidencie a úplnosť
 - sémantická kontrola, dolovanie dát
- Absencia stratégie tvorby dát
 - kontrola kvality dát

Stav

- Vysoká variabilita dát vedie k zložitým štandardom pre ich prenos a uloženie
- Morálna zastanosť tkvie v nekonceptčnosti doplnkov, vysokej voľnosti, čiastočnej nejednoznačnosti a spornosti
 - MARC formáty a interoperabilita (vnútorná, vonkajšia)
- Oddelené štandardy pre formu (MARC) a sémantiku (AACR2, RDA) chaosu výdatne pomáhajú
- Absencia dostatočných centrálnych zdrojov dát celý negatívny obraz reality len umocňuje

Riešenia

- Používať štandardy len na účel na ktorý boli primárne určené
 - ISO2709 a XML len na kódovanie
 - *MARC* len na výmenu štruktúrovaných dát
- Vykonať poctivú analýzu minimálneho a odporúčaného záznamu pre jednotlivé druhy evidovaných objektov
 - Identifikovať entity a ich možné vzťahy
 - Dôsledne selektovať kategorické dáta a dáta voľné

Všetko pozitivne zachovať

- Postaviť formát inšpirovaný FRBR modelom
- Použiť priemyselné štandardy (XML, XSD)
- Trvať na jednoznačnosti dát a znovupoužitelnosti
 - Každý jednoznačne identifikovateľný objekt by mal byť entitou (biblio, osoba, projekt, ..)
 - Vzťahy medzi entitami nemôžu byť len existenčné, ale musia mať možnosť kategorického popisu – odstraňovať dátové duality (biblio – autor)

Problémy modelu

- Mnoho druhov entít databázu zdanlivo triešti
- Znovupoužitelnosť dát vyžaduje viac disciplíny, kontroly a jasnú stratégiu od počiatku
- Vytváranie väzieb medzi entitami môže komplikovať proces tvorby dát
 - odpor k tvorbe autorít (na 10% neprekonaný)
 - odpor k tvorbe heslárov (na 40% neprekonaný)
 - odpor k tvorbe obsahov (na 70% neprekonaný)

Dynamické užívateľské rozhranie

- Dynamické a navádzajúce prostredie
- Našepkávače a kontextové výbery
- Systém zameriavania pozornosti pre minimalizáciu počtu prvkov zobrazených v jednom okamihu na obrazovke
- Integrovaná plná sémantická kontrola pre každú entitu, väzbu a závislosť

Je potrebné vidieť a zažiť !!!

Aktuálne implementácie

- existujúce
 - www.scidap.sk - Scidap (bez prihlásenia len prezentačné funkcie)
- pripravované
 - CREPČ₂
 - CREUČ₂ (2016)
 - Infogate2 (vid'. firemná prezentácia)
 - CRZP₂ (jún 2015)

Pre prístup do demoverzie ma kontaktujte!

Ďakujem za pozornosť

Mgr. Ján Grman, PhD.

grman@svop.sk

www.svop.sk